Office de la propriété intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An Agency of Industry Canada

PCT/CA2004/001448
08 SEPTEMBER 2004(08.09.2004)

REC'D 26 OCT 2004
WIPO PCT

Bureau canadien des brevets
Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

Canadian Patent Office
Certification

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,463,301**, on April 22, 2004, by **ATANAS ILIEV LALEV**, for "Method of Detecting Protein Interactions - Affinity Purification with Switching the Ligand".

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Tracey Faulkner
Agent certificateur/Certifying Officer

September 8, 2004
Date

Canada

(CIPO 68)
31-03-04

OPIC CIPO

Method for detecting protein interactions – affinity purification with switching the ligand (APSL) and uses thereof

Atanas Iliev Lalev and Jack Frederick Greenblatt – Canada

Abstract: The invention is a method for isolating and/or analyzing proteins and/or other biomolecules by separating them from an immobilized protein of interest that has been associated with them in vivo.

Field and background of the invention

[0001] An important scientific pursuit in the post-genomic era is to determine the components of the multiprotein complexes and to identify the protein-protein interactions in a proteome. The proteome consists of stable protein complexes and single (free) proteins but virtually every cellular process is mediated by and/or consists in transient protein interactions, e.g. protein-protein interactions, DNA-protein interactions or RNA-protein interactions. The detection and analysis of the transient interactions is a major challenge in biology and proteomics.

[0002] A preferred method for protein purification and detecting protein interactions is creating fusion proteins by using recombinant DNA techniques. The fusion protein contains a protein of interest and an affinity tag capable of binding selectively to an affinity matrix. After expressing the fusion protein in exogenous or endogenous organism, it is immobilized on an affinity matrix and the unbound substances are removed. Subsequently, the fusion protein is released from the affinity matrix. The method is used in order to purify the affinity tagged protein or, more recently, the affinity tagged protein and the proteins that associate with it. The advent of sensitive mass spectrometric methods for protein identification and improved affinity purification methods made possible the direct identification of protein complexes on a proteome-wide scale (Ho Y. et al Nature 2002, Gavin, A.C. et al. Nature 2002). A review by R. Aebersold and M. Mann, Nature, March 2003, Vol. 422, "Mass spectrometry-based proteomics" describes in more detail the state of the field.

[0003] More than 300 stable complexes have been identified in the model organism Saccharomyces cerevisiae and extensive proteome maps have been built. However, it became clear that the number of transient protein-protein interactions detected by affinity purification coupled with mass spectrometry is smaller (often by an order of magnitude) comparing to the number of transient protein-protein interactions detected by genetic based methods and library based methods. This is due to the fact that when the fusion protein forms transient complexes with other proteins, the latter are isolated in substoichiometric amounts and most often are not detected.

[0004] When the fusion protein and the interacting proteins are isolated via an affinity tag, the bulk of the fusion protein is immobilized on the affinity matrix, whereas only that fraction of the interacting proteins which has been bound to the fusion protein is immobilized. When the fusion protein is released from the affinity matrix, the ratio between the proteins in the eluate remains the same and, as a result, the highly abundant fusion protein suppresses the identification and analysis of the substoichiometrically interacting proteins. The detection of transient protein-protein interactions (e.g. interactions between stable protein complexes or between two single proteins or between a stable protein complex and a single protein) is still very difficult and in most cases impossible. Examples of transient protein-protein interactions are protein substrate – modifying enzyme such as protein substrate - protein kinase, numerous signal transduction events, DNA polymerase – replication factors, RNA polymerase – transcription factors.

[0005] Unless otherwise indicated, the terms "high abundance", "low abundance", "high amount" and "low amount" designate the amounts of the proteins when they are immobilized on the affinity matrix but not their amounts in the organism.
The definition of a transient protein complex as given by S. Jones and J. Thornton, Proc. Natl. Acad. Sci. USA, Vol. 93, 1996 is used – a complex with components that exist as both complexes and free proteins. In the description of the invention, the term "interacting protein" refers to a protein that forms a transient complex with the protein of interest and it is not to be confused with the subunits of a permanent protein complex. "Associated protein" refers to proteins that form both stable and transient complexes with the protein of interest. "Protein of interest" designates the protein whose partners are sought.

[0006] The following terms, synonyms and abbreviations are used in the application: protein interaction = interaction involving protein (e.g. protein-protein, DNA-protein, RNA-protein); protein of interest = protein whose partners are sought; fusion protein = affinity tagged protein = tagged protein; ligand coated solid support = affinity matrix; weak protein interaction = transient protein interaction; stable protein complex = permanent protein complex = core complex; high salt eluate = eluate that has been obtained by increasing the ionic strength of the medium surrounding the immobilized complexes; SpA-tag = affinity tag consisting of one or more IgG binding domains genetically derived from Protein A from Staphylococcus aureus; SpG-tag = affinity tag consisting of one or more IgG binding domains genetically derived from Streptococcal protein G (strain G148); dual affinity tag = two different affinity tags fused together; interacting protein = protein interactor = protein that binds directly or indirectly to the protein of interest and forms a transient complex; cellular lysate = protein extract; unbound substances = substances that are not immobilized directly or indirectly on the affinity matrix; protein complex = biological complex containing one or more protein components, pre-drug = drug candidate = chemical that is tested in order to determine whether it selectively affects a particular complex associated with a disease; SDS-PAGE = polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulphate; M = molar (concentration); mM = millimolar; mcM = micromolar; ml = milliliter, mcl = microliter; rpm = revolutions per minute (during centrifugation); MBP = maltose binding peptide; SBP - streptavidin binding peptide; WB = washing buffer; MALDI-TOF MS = matrix assisted laser desorption ionization / time of flight mass spectrometry, contaminant protein = protein that binds nonspecifically to the protein of interest or to any protein or to the affinity matrix.
In the description of the invention, the term protein complex means a biological complex containing one or more protein components. Other components can be nucleic acids or lipids. Proteins can be modified (e.g. glycoproteins or lipoproteins). Washing buffer means the buffer that is used to remove the unbound substances. Elution buffer means the buffer that is used to separate the interacting proteins from the immobilized protein of interest. Most often, the protein of interest is part of a fusion protein and the elution buffer is washing buffer with an increased concentration of KCl or other salt and/or other substances.

[0007] The following example with RNA polymerase II and the transcription factors illustrates the problem with detection of transient protein-protein interactions. Although the invention is

illustrated here by SpA-tagged subunits of RNA Polymerase II from Saccharomyces cerevisiae immobilized on IgG-beads, the example demonstrates a general problem with detecting any transient protein interaction by using any affinity tagged protein from any organism.

[0008] Eukaryotic RNA polymerase II is a permanent complex of twelve subunits and it forms transient complexes with different transcription factors. When a SpA-tagged subunit of RNA Polymerase II core complex from Saccharomyces cerevisiae is expressed under physiological conditions, it associates permanently with other subunits of the core complex and, as a part of the core complex, it interacts directly or indirectly with a variety of transcription factors by forming transient (weak) complexes. After preparation of a protein lysate and immobilization of the SpA-tagged subunit on IgG-beads, the other subunits of the core complex are present in approximately equal molar amounts but the transcription factors are present in substoichiometric amounts.

[0009] Several factors contribute to the unequal stoichiometries: (a) during the preparation of the cellular lysate and subsequent purification steps, transient complexes containing the RNA Polymerase II core complex dissociate in different ways and each core complex remains associated with a different set of transcription factors, (b) during the synthesis of mRNA, the RNA Polymerase II core complex associates with different proteins at different stages of the transcription, e.g. initiating factors, mediator complex, elongation factors, termination factors, (c) during the transcription of different genes, RNA Polymerase II core complex associates with different sets of gene specific transcription factors, (d) a certain amount of the RNA Polymerase II core enzyme is not involved in transcription and is not associated with other proteins.

[0010] As a result, the amounts of the subunits of the core complex on the affinity matrix are disproportionably higher than the amounts of the interacting proteins. The isolation of the complexes by releasing the SpA-tagged subunit from the IgG-sepharose matrix and their analysis by mass spectrometry leads to identification of the subunits of the core complex, subunits of the general transcription factor TFIIF and, in some cases, Spt5, Taf10, Ceg1, Ess1, Kin28 and Srb6. However, more than 50 proteins that interact physically with the RNA polymerase II core enzyme are reported in the literature (Myer V.E. and Young R.A., J.Biol.Chem 1998).

[0011] Even when the isolated RNA Polymerase II complexes are fractionated by electrophoresis, some of the twelve subunits of the core complex (most often Rpb1 and Rpb2) and their different modification forms and degradation products are present in large amounts in many bands and they suppress the identification and/or analysis of low abundance proteins by mass spectrometry. In this case, the mass spectrum contains very large peaks, resulting from the high abundance proteins (i.e. subunits of Polymerase II core complex), and small peaks, resulting from the low abundance interacting proteins (i.e. transcription factors). Most often, the small peaks are not even detected and the analysis of the spectrum leads to identification of only the major protein(s). This is due to the following problems in mass spectrometry: (i) The resolution problem in mass spectrometry instruments – if a large peak and small peak are separated by only a few mass units (i.e. their molecular weight differs by only 3-4 Da), the detection of the latter is impossible, (ii) the dynamic range problem in mass spectrometry (high background noise problem) - the signals from the abundant peptides suppress the signal from the minor proteins, even if they are separated by more than 3-4 mu. The background of the mass spectrum is due to the presence of large numbers of ions with low intensity. They are mostly multiple-charge ions and fragment ions from the major peptides and they are not detected as signals but as a background noise in the mass spectrum. In addition, adducts that are bound to the abundant peptides increase the background (chemical noise). (iii) Ionization of the MALDI-TOF samples (co-crystallized peptides and a UV-matrix) is most optimal when the molar ratio between the matrix and the peptides is between 2,000:1 and 20,000:1. Hence, it is impossible to achieve a good ionization, and consequentially a good spectrum, for both high abundance and low abundance peptides, when the ratio between their amounts is 100:1 or 1000:1 or higher.
The example demonstrates the problems that are encountered when transient interactions are investigated by affinity purification coupled with mass spectrometry. The problem is even more severe in gel-less protein identification approaches (e.g. LS/MS/MS) because of the absence of initial separation step (PAGE) of the isolated proteins.

[0012] Besides the dynamic range problem in mass spectrometry, other problems arise when low abundance proteins or other substances are analyzed together with high abundance ones: (a) cross-reactivity problem in immunoassays, (b) overloaded chromatography columns or overloaded protein gels, and as a result, poor resolution, (c) ambiguous results in enzymatic assays.

It is therefore a goal of the present invention, Affinity Purification with Switching the Ligand – APSL, to solve these problems by separating the substoichiometrically interacting proteins and/or other biomolecules from the high abundance protein of interest.

Description of the invention

[0013] The present invention is a method for isolating and analyzing proteins and/or other biomolecules that form transient complexes with a protein of interest in vivo. Although affinity tagged proteins were used in the examples (see gel images) and the interacting proteins were identified by MALDI-TOF MS, the scope of the invention is not limited to the field of spectrometry-based proteomics. The essence of the invention is isolating and detecting proteins and/or other biomolecules that interact in vivo with a protein of interest by the following procedure: (a) introducing into an organism or a cell a heterologous nucleic acid encoding an affinity tag fused to a protein of interest or part of it, (b) expressing the fusion protein under physiological conditions that enable formation of stable and transient complexes and obtaining a cellular lysate or other biological fluid containing complexes that include the protein of interest, (c) selectively immobilizing the protein of interest on an affinity matrix and removing the unbound substances, (b) isolating and/or analyzing the interacting proteins and/or other biomolecules by separating them from the immobilized protein of interest.
Besides the protein-protein interactions, the invention is suitable for detecting other interactions that involve proteins, e.g. DNA-protein interactions.
Another aspect of the invention is a method for drug discovery. A chemical or a biomolecule can be identified as a drug or a pre-drug by its ability to affect selectively a particular protein interaction that is associated with a disease.

[0014] The main distinctive feature of the invention is that the fusion protein or tagged subunit of protein complex remains immobilized on the affinity matrix but the interacting proteins and/or other biomolecules are separated into a liquid phase and analyzed separately. This separation facilitates the identification and analysis of the interacting proteins and/or other substances that are associated with the protein or protein complex of interest in substoichiometric amounts. For example, the identification of the low abundance proteins by mass spectrometry is facilitated if

they are analyzed separately from high abundance ones because of the elimination of a major problem in mass spectrometry – the dynamic range problem.

[0015] As described above, naturally assembled RNA Polymerase II complexes can be immobilized on IgG-beads via a SpA-tagged subunit of the core complex. We have found that when the immobilized complexes are treated with buffer containing 0.4 M KCl, the transcription factors are separated into the liquid phase but the core complex remains immobilized on the IgG beads. In addition, the identification and analysis of the transcription factors by mass spectrometry was greatly improved by the separation.

[0016] The invention is based on the finding that increasing the ionic strength does not disrupt either the SpA:IgG bond or the bonds between the subunits of permanent protein complexes (e.g. the RNA Polymerase II core complex) but leads to rapid and almost complete dissociation of transient protein complexes (e.g. transcription factors). Since IgG:ProteinA binding (Fc-reactivity) has a mainly hydrophobic nature, the same must be the case for binding between the subunits of stable complexes (i.e. RNA Polymerase II). Transient interactions between proteins are due in large part to electrostatic interactions, as demonstrated in the examples.

[0017] Increasing the ionic strength destabilizes the electrostatic attraction between charged groups (salt bonds) because of the Debye-Huckel screening and according to Coulomb's law: $F = q_1q_2/Dr^2$, where: F – strength of the electrostatic force, q - charge of the ion; D - effective dielectric constant of the media; r - distance between the ions. On the other hand, increasing the salt concentration stabilizes the hydrophobic bonds by competing out the water molecules from the non-polar patches of the protein surface so that they associate even stronger by hydrophobic interaction. This increase in the entropy of water molecules, as the number of them solvating hydrophobic surfaces decreases, is the main driving force for protein precipitation by "salting out" techniques (e.g. ammonium sulphate precipitation).

[0018] The fact that more than 60 interacting proteins can be dissociated from the core complex with 12 subunits by weakening the electrostatic bonds (i.e. by increasing the ionic strength of the medium) allows the formulation of a general proteomic rule: "Permanent protein complexes are held together mainly by hydrophobic forces and transient protein complexes are held together in

large part by electrostatic attractions (mainly salt bonds between the charged amino acids)". The salt bonds between the subunits of permanent protein complexes are not affected by increasing the ionic strength because the ions cannot reach into the hydrophobic interface between the subunits and weaken them. The salt bonds that are formed during a transient protein-protein interaction are not shielded by the hydrophobic interface and they are easily disrupted by increasing the ionic strength of the medium. The other electrostatic attractions, which are much weaker, i.e. London forces, dipole-dipole, charge-dipole, are also weakened by increasing the ionic strength. (Of course, some minor exceptions from the above proteomic rule are possible)

[0019] Statistical analysis by Jones, S and Thornton, J., PNAS 1996 showed that the contact surfaces (interfaces) of transient protein complexes contain more hydrophilic residues than the contact surfaces of permanent protein complexes. Our results demonstrate that among all hydrophilic residues, the most important for formation of transient protein complexes are electrostatically charged ones - Lysine (positive), Arginine (positive), Aspartate (negative), Glutamate (negative) and partially Histidine (50% positively charged at pH 6.0), that can form salt bonds or other electrostatic bonds. Another group of amino acids that are involved in transient protein-protein interactions are the ones that are electrostatically charged as a result of post-translational modifications, e.g. phosphorylation, acetylation.
It has already been proposed, based on a statistical analysis (Archakov, A. et al. Proteomics, 2003), that formation of permanent complexes resembles and even might be a continuation of protein core folding, and henceforth, is due to the hydrophobic force. Our results support this point of view and further suggest that stable protein complexes may have a common hydrophobic interior that is not accessible to the ions. The same authors make a correct statement that the nature of transient protein-protein interactions was not clear at the time of the publication. We have found that transient protein-protein interactions are predominantly electrostatic. Moreover, since among all electrostatic forces charge-charge interactions (Coulomb force) are the strongest ones, they play a major role in transient protein interactions.

[0020] Three interesting features of the Coulomb force explain our finding that they are responsible for the formation of the majority of the transient protein complexes: (I) Coulomb forces are long-range forces and they still play a role at 1 nanometer distance. All other forces are short-range forces, e.g. London forces occur when the molecules are close enough to induce

electrostatic dipoles in one another. (II) Coulomb forces are the strongest ones among noncovalent forces (up to -80 kcal/mol for salt bonds). The other noncovalent bonds are much weaker. Hydrogen bonds are -3 to -6 kcal/mol, Van der Waals bonds are -0.5 to -1 kcal/mol and hydrophobic bonds are -0.5 to -3 kcal/mol. (III) Coulomb forces can be either attractive or repulsive, depending on the nature of the charges involved, i.e. attractive between positive and negative charges, and repulsive between two positive charges or between two negative charges. Most probably, the formation and disintegration of transient protein complexes is due to alternation of electrostatic attraction and electrostatic repulsion between the electrostatically charged amino acids. The switch could reflect the accomplishment of the function that requires formation of the transient protein complex and the three most plausible causes for the switch are: (a) change of orientation and/or position of the interacting proteins relative to one another, and/or (b) change of the conformation of one or both of the interacting proteins and/or (c) adding or removing an electrostatic charge as a result of post translational modification, e.g. acetyl group to Lysine or phosphate group to Serine, Tyrosine and Threonine.

[0021] Whatever the exact mechanism of transient protein-protein interactions might be, it is clear that electrostatic forces are of primary importance. The invention is designed to detect transient protein-protein interactions by weakening the electrostatic forces and separating the interacting proteins from the immobilized protein of interest.

[0022] The invention has several significant advantages comparing to other methods for protein purification and/or detecting protein interactions.
(I) The method of the invention does not require a purified protein of interest for the preparation of affinity column,
(II) The immobilized protein or protein complex is in its natural modification state (e.g. phosphorylation state), which is very often necessary for proper formation of protein complexes,
(III) Since many proteins exist and interact as components of protein complexes, but not as single proteins, the invention is suitable for detecting transient protein interactions that involve multicomponent protein complexes.
(IV) The method of the invention detects proteins that have been associated with the fusion protein in vivo. Protein affinity chromatography according to Formosa T., Methods Enzymol.

1991, and GST pull downs are based on in vitro (de novo) association of the proteins with the immobilized protein and that makes possible many nonspecific interactions.
(V) The main advantage of the invention over protein purification via one or more affinity tags is the separation of the substoichiometrically interacting proteins from the tagged protein or protein complex and, as a result, greatly facilitated analysis and identification.

[0023] The invention can be carried out with any organism or a cell line from an organism, provided that the sequence of the gene of interest is known. A heterologous nucleic acid encoding a selectable marker and an affinity tag fused in-frame with the protein of interest (or a part of the protein of interest) is constructed and administered into the said organism or a cell from an organism. After introducing the heterologous nucleic acids into the cell, the fusion protein can be expressed extrachromosomally and, optionally, the chromosomal copies of the gene can be silenced. Preferably, chromosomal copies of the gene of interest can be replaced by homologous recombination. Standard DNA recombinant techniques and yeast manipulation techniques are described in Sambrook, Fritsch, Maniatis, 1982 Molecular Cloning, Cold Spring Harbor Laboratory Press, and Current Protocols in Molecular Biology, John Wiley and Sons, Inc, New York, 1994.

[0024] The heterologous nucleic acid expressing the fusion protein contains sequences ensuring its proper transcription, processing and translation. The fusion protein should be expressed under physiological conditions enabling its associations with other proteins or biomolecules. The fusion protein may contain the entire protein of interest or a part of it.

[0025] The invention involves usage of chromatography columns and fusion proteins.
Figure 1 shows exploded views of a chromatography column packed with ligand coated solid support: 1 – chromatography column packed with affinity beads, 2 – an expanded view of several individual beads, 3 – an expanded view of an individual bead of affinity matrix, 4 – solid support (e.g. agarose or sepharose), 5 – ligand (e.g. IgG or glutathione) that binds selectively to affinity tag, 6 – linker between the solid support and the ligand.

[0026] Figure 2 illustrates the two essential steps (bold vertical arrows) of the invention Affinity Purification with Switching the Ligand (APSL). Legend: 7 – linker between the protein of interest

and the affinity tag, 8 – affinity tag that is fused to a protein of interest; 9 – protein of interest or subunit of a permanent protein complex that is affinity tagged (i.e. fusion protein); 10, 11, 12, 13, 14, 15 – proteins that interact directly with the fusion protein or protein complex; 19 - protein that interacts indirectly with the fusion protein or protein complex; 16 – the fusion protein or a protein complex and the interacting proteins are immobilized on the affinity matrix; 17 – the fusion protein or protein complex remains immobilized on the affinity matrix after the elution; 18 – the interacting proteins are separated into the liquid phase after the elution (the eluate is indicated with a dotted line); 20 - Step One: preparation of cellular lysate or other biological fluid from an organism or cell line containing a heterologous nucleic acid expressing an affinity tagged protein and immobilizing the affinity tagged protein on the affinity matrix by the tag; removing the proteins and other substances that are not bound directly or indirectly to the fusion protein; 21 - Step Two: separating the interacting proteins from the immobilized fusion protein by elution with an agent that does not disrupt the binding between the affinity matrix and the immobilized fusion protein or subunit of a permanent protein complex; only proteins and other substances that bind directly or indirectly to the immobilized protein are eluted. The bonds between some of the affinity tags and their respective ligands are mainly hydrophobic and it is therefore reasonable to increase the ionic strength of the medium in Step Two in order to weaken the electrostatic attractions by Debye-Huckel screening.

[0027] The essence of the invention does not consist in using a particular affinity tag or a particular affinity matrix. Elsewhere in the application, several examples of carrying out the invention with a protein of interest devoid of affinity tag are described. The essence of the invention consists in isolating and detecting proteins by disrupting in vitro the associations with the protein of interest that have been formed in vivo.

[0028] The novelty of the invention consists in the separation of the interacting proteins and/or other substances from the immobilized fusion protein. The name of the method reflects the fact that the fusion protein serves as a ligand for the interacting proteins in vivo and binds to another ligand in vitro. The term ligand is used here as a molecule that binds another molecule (Latin: ligare = to bind).

[0029] Although in most cases the invention includes usage of affinity tagged proteins, the invention not only differs from the methods for purification via affinity tag but is even quite opposite to them, i.e. the greatest care must be taken to avoid the separation of the fusion protein from the affinity matrix during the elution. The novelty of the invention is apparent by comparison with the other methods for affinity purification as described by Bauer A. and Kuster B. in Eur. J. Biochem. Feb 2003, or, by comparison with the other methods for detection and analysis of protein interactions as described in a review by Phizicky E. and Fields S. in Microbiological Reviews, Mar. 1995, Vol. 95, No. 1.

[0030] Moreover, all basic methods for protein purification (i.e. affinity chromatography, hydrophobic interactions, ion exchange, reversed phase) include separation of the protein that has been bound to the solid phase. The invention offers a convenient way for separating substoichiometrically associated proteins and/or other biomolecules from the immobilized protein and constitutes a novel basic method for purifying proteins and detecting protein interactions.

[0031] The biomolecule of interest can bind selectively to the affinity matrix by itself or via an affinity tag which facilitates the selective immobilization of the biomolecule of interest. Many methods allow for selective binding of proteins or other biomolecules devoid of an affinity tag to an affinity matrix. These methods are commonly referred as affinity chromatography and all of them include binding and separation of the biomolecule of interest from the affinity matrix.

[0032] The invention can also be described as immobilizing selectively a complex containing the biomolecule of interest on a solid phase by binding between the biomolecule of interest and the solid phase and, after removing the unbound substances, separating the associated biomolecules from the immobilized biomolecule of interest. The biomolecule of interest should bind to the solid phase by dominantly one type of force (preferably, hydrophobic force) and elution of the interacting biomolecules should be performed by weakening another force. As shown in the examples of Rpb1 fused to a SpA-tag (see Figure 2,3 and 4), the biomolecule of interest (i.e. the fusion protein) binds to the affinity matrix (i.e. IgG beads) by hydrophobic force (i.e. Fc reactivity) and the interacting proteins (i.e. transcription factors) are separated by weakening the electrostatic bonds (i.e. increasing the ionic strength of the medium).

[0033] A nucleoprotein complex can be immobilized by binding between its nucleic acid component and an affinity matrix coated with complementary sequence. The nucleic acid component of the complex can be genetically engineered so that it contains a Poly(G) and the complex can be immobilized on Poly(C) matrix. The associated proteins can be eluted with increasing the ionic strength of the medium. In this case, the biomolecule of interest is bound to the affinity matrix by hydrogen bonds the interacting proteins are separated by increasing the ionic strength of the medium.

[0034] Separation of the interacting protein(s) from the immobilized protein or protein complex can be achieved by enzymatic treatment (e.g. treatment that removes or adds an electrostatic charge as a result of acetylation or deacetylation or phosphorylation or dephosphorylation). In this case, essential information can be obtained about the nature of a particular protein-protein interaction.

[0035] The method is especially suitable for detecting proteins that interact with a multicomponent protein complex. It is best if the elution is performed under such conditions that all subunits of the permanent protein complex remain immobilized because a significant leakage into the liquid phase of any subunit, not only the tagged one, will suppress the identification and analysis of the substoichiometrically interacting proteins.

The following Examples of carrying out the invention are illustrative but not limiting the scope of the invention.

[0036] Heterologous DNA encoding a selectable marker and a SpA-tag fused in-frame to the carboxy-terminus of Rpb1 was constructed and administered into Saccharomyces cerevisiae cells. The resulting strain expresses a fusion protein containing Rpb1 and a SpA-tag containing four IgG-binding domains. After selecting the transformants, the proper expression of the fusion protein was checked by Western blotting. Cell cultures were grown in YPD medium at 30°C and constant vigorous shaking and cells were collected at middle or late logarithmic phase. 9 liters culture was grown in YPD medium to $OD_{600} = 1.2$. The following protocol was used:

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations were performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine, 2x phosphatase inhibitors.
5. Centrifugation at 16,000 rpm for 30 minutes and transferring the clarified lysate to a clean tube.
6. Adding 800 microliters of IgG beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the IgG beads to four parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 3ml WB containing 0.4M KCl – WB (0.4M)
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB (0.4M) and combining the eluates.
12. Distributing the eluate to eppendorf tubes and adding 1/5 volume of 100% TCA.
13. Incubating on ice for 1 hour; centrifugation for 1 hour at 14,000 rpm at 0°C to 5°C.
14. Decanting the supernatant and adding 1 ml 90% acetone. Vortexing for 30 seconds and centrifuging for 10 minutes at 14,000 rpm. Decanting the supernatant.
15. Drying with a SpeedVac for 15 seconds.
16. Protein gel electrophoresis (SDS-PAGE - 10% gel, 20 cm long) and silver staining.
A gel image of high salt eluate from immobilized RNA Polymerase II from Saccharomyces cerevisiae is shown in Figure 3. The digits and arrows at the right side of the gel indicate the identified proteins. Legend: 31 – Spt6, 32 – Rgr1, 33 – Spt6 + Spt5, 34– Tfg1 + Spt5, 35– Set3, 36 – Cet1 + Med1, 37 – Pob3, 38 - Npl3, 39 – Tfg2 + Iws1, 40 – Cdc73, 41 – TFIIS, 42 – Ess1.

[0037] Protein identifications by mass fingerprinting were made essentially as described in Shevchenko A et.al. Anal. Chem., 1996, and Van Montfort, B.A. et.al., J. Mass Spectrom., 2002. After the SDS-PAGE and silver staining the gel lanes were cut across the whole length (including the bands stained most weakly with silver and "blank" regions) and the proteins in the gel pieces were reduced by DTT treatment and alkylated by iodoacetamide treatment. After trypsin digestion, the resulting peptides were purified and concentrated by reverse phase chromatography on C18-matrix. The peptides were co-crystallized with CHCA (a-cyano-4-hydroxycinnamic acid) ionization matrix and subjected to MALDI-TOF MS. After obtaining the MALDI-TOF MS spectra, the proteins were identified by matching the list of experimentally obtained peptide masses with computer-generated lists of peptide masses for every predicted protein in a Saccharomyces cerevisiae databanks. When contaminant proteins were identified in the sample, the corresponding peptide masses were subtracted and the search was performed again with the remaining masses. When a band is not indicated by an arrow, it means that it contains alternative translation products of proteins that are indicated elsewhere or their degradation products or contaminant proteins. The proteins identified with high confidence are listed in the text (the legends) and indicated in the Figures.

[0038] A gel image of high salt eluate from immobilized RNA Polymerase II from Saccharomyces cerevisiae is shown in Figure 4. The fusion protein contains Rpb1 and a SpA tag containing two IgG binding domains. 18 liters culture was grown to $OD_{600} = 0.7$ and the separation of the interacting proteins was performed with washing buffer containing 0.4 M KCl - WB(0.4M). The following protocol was used:

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations were performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine, 2x phosphatase inhibitors.

5. Centrifugation at 16,000 rpm for 30 minutes and transferring the clarified lysate to a clean tube.
6. Adding 800 microliters of IgG beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the IgG beads to four parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 3ml WB containing 0.4M KCl – WB (0.4M)
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB (0.4M) and combining the eluates.
12. Distributing the eluate to eppendorf tubes and adding 1/5 volume of 100% TCA.
13. Incubating on ice for 1 hour; centrifugation for 1 hour – 14,000 rpm at 0°C to 5°C.
14. Decanting the supernatant and adding 1 ml 90% acetone. Vortexing for 30 seconds and centrifuging for 10 minutes at 14,000 rpm. Decanting the supernatant.
15. Drying with a SpeedVac for 15 seconds.
16. Protein gel electrophoresis (SDS-PAGE - 10% gel, 20 cm long) and silver staining.

Legend: 51 – Spt5 + Spt6, 52 - Spt5 + Spt6, 53 - Spt5 + Spt6, 54 - Ctr9 + Spt5 + Spt6, 55 - Spt5 + Ydl145 (Cop1) + Spt6, 56 - Tfg1 + Spt5 + Spt6 + Spt16, 57 - Spt5 + Tfg1 + Sec21, 58 - Spt5 + Tfg1, 59 - Set2 + Spt5 + Rtf1 + Ydl145 (Cop1), 60 - Cet1, 61 - Tfg2 + Iws1, 62 - Ceg1, 63 - Cdc73, 64 – TFIIS, 65 - Rtt103, 66 - Tfg3, 67 - Yhr121, 68 - Ess1.

[0039] Besides the separation of the interacting proteins and other biomolecules from the immobilized fusion protein, increasing the ionic strength causes separation of residual contaminant proteins from the affinity matrix. One way to avoid false signals (i.e. eliminate the background from the contaminants) is to perform the method of the invention with cellular lysate from an organism devoid of affinity tagged protein and identify the proteins in the eluate as contaminants. Alternatively, the proteins that are isolated by performing the method of the invention with several biologically unrelated proteins can be compared and the common ones can be identified as contaminants. The contaminants are not indicated in the Examples.

[0040] An example of identifying proteins that bind nonspecifically to IgG beads and/or to other proteins (contaminants) is shown in Figure 5. Protein extract from Saccharomyces cerevisiae lacking an affinity tagged protein was incubated with IgG beads, the unbound proteins were removed and proteins that bind to IgG-sepharose and/or to other proteins nonspecifically were eluted with high salt in the washing buffer. The same protocol as in Figure 4 was used. Legend: 70 – Acc1+ Fas1 + Fas2, 71 – Lys2 + Ydr098, 72 – Mdn1 + Tra1, 73 – Sec27 + Ade3, 74 – Prt1 + Lys2, 75 – Sse1 + Lsg1, 76 – Ssb1 + Ssa2 + Pab1 + Ded1, 77 – Hsp60 + Dhh1, 78 – Gcd11 + Ded1 + Nop12 + Nog2, 79 – Tef2 + Phr1, 80 – Tef2 + L3, 81 – Nop3 + Tef2, 82 - Yor060 + Ydj1, 83 – Stm1 + Rpl4, 84 - Tif34, 85 – Rpp0 + Rpl5, 86 – Tef2 + Nop1, 87 – Rps1 + Rpl2, 88 – Rps3 + Rpl8, 89 – Rps2, 90 – Rps5, 91 – Rps8 + Rpl9, 92 – Rpl10, 93 – Rps20, 94 – Rpl20, 95 – Rps14, 96 – Rps17, 97 – Rpl27 + Rps24. The following contaminants are not indicated in the other figures: Phr1, Tef2, Nop1, Acc1, Nop3, Fks1, Yor060, Fas2, Ydj1, Fun12, Stm1, Aro1, Rpl4, Lys2, Tif34, Ydr098, Rpp0, Rpg1, Rpl5, Rps3, Ade3, Rps1b, Sec27, Rpl2, Prt1, Rpl8, Tef2, Rps2, Lsg1, Rps5, Dhh1, Rpl10, Ded1, Rpl9, Ssz1, Hsp60, Rps8, Gcd11, Rps20, Nop12, Rpl20, Nog2, Pab1, Rps17, Fas1, Ssb1, Ssa2, Sse1, Rpl27, Rps14, Rps24, Hfa1, Zuo1, Asc1, Mdn1, Tra1, Dyn1, Myo2, Ira2.

[0041] After identifying the contaminants for a particular organism, the possibility to assign a contaminant protein as a true interactor, i.e. to produce a false positive interactor, are minimal and the experiments can be performed under less stringent conditions. For example, after immobilizing the complexes on the affinity matrix, the removal of the unbound proteins can be very brief (i.e. washing 200mcl IgG beads with 6 ml washing buffer) and the elution of the interacting proteins can be performed by a higher salt concentration (i.e. 0.7 M potassium acetate). High salt eluates from immobilized RNA Polymerase II complexes from Saccharomyces cerevisiae obtained in three different experiments from three different strains each expressing different SpA-tagged subunit are shown in the following Figures. SpA-tags contain two IgG binding domains. The fusion protein in Figure 6 contains Rpb1, the fusion protein in Figure 7 contains Rpb3, the fusion protein in Figure 8 contains Rpb9. The eluates were resolved by SDS-PAGE and the proteins were identified by MALDI-TOF MS. Legend for Figure 6: 120 - Spt6, 121 – Spt6, 122 - Spt5 + Cyr1 + Clu1, 123 – Tfg1 + Nam7 + Spt16, 124 – Fcp1 + Cdc48 + Sec21, 125 – Taf5 + Top2, 126 – Cet1 + Pcf1 + Prp6, 127 – Pob3, 128 – Rtt103 + Ptc3 + Has1, 129 – Tfg2 + Iws1 + Pap1, 130 – Tfg2 + Iwr1; 131 – Rai1, 132 – Sua7, 133 – TFIIS, 134 –

Ypl253, 135 – Ssn3 + Mrt4, 136 – Tfg3 + Ptc1, 137 – Med7, 138 – Yhr121, 139 – Mbf1 + Taf9 + Yor379, 140 – Ybr262 + Ybr174.
Legend for Figure 7: 150 – Spt6, 151 – Spt5 + Ctr9 + New1, 152 – Spt16 + Tor2, 153 – Tfg1 + Hpr5 + Nam7, 154 – Fcp1 + Cdc48 + Nam7, 155 – Fcp1 + Rad3, 156 - Fcp1 + Taf5 + Pbp1, 157 – Cet1 + Cyr1, 158 – Pob3, 159 – Rtt103 + Ptc3, 160 – TFIIH + Rad23 + Drs1, 161 – Tfg2 + Iws1, 162 – Rai1 + Pob3 + Rgr1 + Ygr086, 163 – Fun11 + Sub1, 164 – TFIIB, 165 – TFIIS + Yol045, 166 – Tfg3, 167 – Din7, 168 – Yhr121, 169 – Nhp2, 170 – Mbf1, 171 – Ykr057.
Legend for Figure 8: 180 – Spt6 + Stt4 + Ydl145 (Cop1), 181 – Spt5 + New1 + Tof1, 182 – Spt5 + Rad3, 183 – Tfg1 + Cdc48 + Rgr1 + Nam7, 184 – Fcp1 + Cdc48 + Rad1 + Sec21, 185 – Fcp1 + Sup35 + Rap1, 186 – Cet1 + Rap1 + Cdc6, 187 – Tfg2 + Iws1, 188 – Ceg1 + Kin1 + Prp8, 189 – Ceg1 + Rgm1, 190 – Ygr086 + Ydr065, 191 – TFIIB, 192 – TFIIS + Tho1 + Bim1, 193 – Tfg3 + Tfa2 (TFIIE), 194 –Set3 + Tfb2 (TFIIH) + Yol022, 195 – Ykr057. Cdc6, Prp2, Prp16, Rif1 and Cft1 were also detected.

[0042] Yet another way to reduce the possibility of producing false positives is to reduce the background from the contaminant proteins by performing an additional purification step before the separation of the interacting proteins from the fusion protein. Preferably, the additional step is an affinity purification step and in this case, two different affinity tags can be fused to the protein of interest (i.e. protein is fused to a dual tag) or to two subunits of a protein complex, and, after the second immobilization, the interacting proteins are separated from the immobilized fusion protein. This variant of the invention consists of two methods, i.e. protein purification by an affinity tag followed by the basic method of the invention. Currently, the most popular dual tags are the TAP (tandem affinity purification) tags containing a calmodulin binding domain and IgG binding domain separated by a Tev cleavage site. The tag used in the following experiments is a modified version of the TAP-tag (Ghaemmaghami, S., et.al. Nature, 2003).

[0043] A gel image of high salt eluate from immobilized TFIIF from Saccharomyces cerevisiae is shown in Figure 9. The carboxy-terminus of the chromosomal copy of the gene encoding Tfg1 is fused in-frame with a TAP-tag. 18 liters Saccharomyces cerevisiae culture was grown in YPD medium to $OD_{600} = 0.7$. The complex was purified on IgG beads and further immobilized on calmodulin beads. The elution was performed with 0.3 M KCl in the buffer. Legend: 201 - Rpb1, 202 - Rpb1+Rpb2, 203 - Fcp1 + Rpb2 + Rpb1, 204 - Rpb1 + Rpb2 + Fcp1 , 205 - Glo3, 206 -

Rpb3, 207 – TFIIS, 208 - Mpd2, 209 - Rpb4, 210 - Yhr121 + Rpb5, 211 - Rpb6, 212 - Rpb7, 213 - Mbf1 + Rpb8, 214 – Rpb9. The following protocol was used:

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations were performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine, 2x phosphatase inhibitors.
5. Centrifugation at 16,000 rpm for 30 minutes and transferring the supernatant to a clean tube.
6. Adding 800 microliters of IgG beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the IgG beads to three parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Washing with 400 microliters WB containing 2mM DTT.
10. Transferring the beads with 2ml WB containing 2mM DTT to two eppendorf tubes and adding 500 units Tev protease. Rotation at 0°C to 5°C for 2 hours.
11. Transferring the beads and the buffer back to 10ml chromatography column and collecting the eluate. Washing the beads with two more column volumes of WB containing 2mM DTT and combining the eluates in a 5 ml tube.
12. Adding of 500 microliters calmodulin beads and 5mM $CaCl_2$ and 3 ml WBC (washing buffer for calmodulin beads). WBC - 100 mM KCl, 10 mM HEPES pH 7.4, 5 mM DTT, 5 mM $CaCl_2$, 1 mM immidazol, 5% glycerol, 0.1% TritonX100. 1 hour rotation at 0°C to 5°C.
13. Transferring the beads to 10ml chromatography column and washing with 5ml WBC.
14. Transferring the beads to a 5ml tube and adding 3ml WBC containing 0.3M KCl - WBC(0.3M). Rotating the tube for 10 minutes.

15. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with 2 column volumes of WBC (0.3M) and combining the eluates.
16. Distributing the eluate to eppendorf tubes and adding 1/5 volume of 100% TCA.
17. Incubating on ice for 1 hour; centrifugation for 1 hour – 14,000 rpm at 0°C.
18. Decanting the supernatant and adding 1 ml 90% acetone. Vortexing for 30 seconds and centrifuging for 10 minutes at 14,000 rpm. Decanting the supernatant.
19. Drying with a SpeedVac for 15 seconds.
20. Protein gel electrophoresis (SDS-PAGE - 10% gel, 20 cm long) and silver staining.

[0044] A gel image of high salt eluate from immobilized RNA Polymerase II from Saccharomyces cerevisiae is shown in Figure 10. The carboxy-terminus of the chromosomal copy of the gene encoding Rpb1 is fused in-frame with a TAP-tag. 18 liters Saccharomyces cerevisiae culture was grown in YPD medium to $OD_{600} = 0.7$. The complexes were purified on IgG beads and further immobilized on calmodulin beads. The elution was performed with 0.5 M KCl in the WBC. A minor release of the fusion protein from the affinity matrix was observed. Legend: 221 - Rpb1, 222 - Spt6, 223 – Spt6, 224 - Rpb2 + Ctr9, 225 – Ctr9 + Spt6, 226 - Tfg1 + Prp22 + Spt16 + Spt6, 227 - Tfg1 + Fcp1 + Spt16 + Nip1, 228 - Fcp1 + Set2 + Rtf1 + Ygl244, 229 - Tfg1 + Fcp1 + Ygr054 + Leo1 + Med1, 230 - Pob3, 231 - Paf1 + Glo3, 232 - Iws1 + Tfg2, 233 – TFg2, 234 - Sua7 (TFIIB homolog), 235 - TFIIS, 236 - Mpd2 , 237 - Tfg3, 238 – Ess1, 239 – Mbf1. The following protocol was used:

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations were performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine, 2x phosphatase inhibitors.
5. Centrifugation at 16,000 rpm for 30 minutes and transferring the supernatant to a clean tube.

6. Adding 800 microliters of IgG beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the IgG beads to three parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Washing with 400 microliters WB containing 2mM DTT.
10. Transferring the beads with 2ml WB containing 2mM DTT to two eppendorf tubes and adding 500 units Tev protease. Rotation at 0°C to 5°C for 2 hours.
11. Transferring the beads back to 10ml chromatography column and collecting the eluates. Washing the beads with two more column volumes of WB containing 2mM DTT and combining the eluates in a 5 ml tube.
12. Adding of 500 microliters calmodulin beads and 5mM $CaCl_2$ and 3 ml WBC. WBC: 100 mM KCl, 10 mM HEPES pH 7.4, 5 mM DTT, 5 mM $CaCl_2$, 1 mM immidazol, 5% glycerol, 0.1% TritonX100. 1 hr rotation. Transferring the beads to 10ml chromatography column and washing with 5ml WBC.
13. Transferring the beads to a 5ml tube and adding 3ml WBC containing 0.5M KCl – WBC(0.5). Rotating the tube for 10 minutes.
14. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WBC(0.5M) and combining the eluates.
15. Distributing the eluate to eppendorf tubes and adding 1/5 volume of 100% TCA.
16. Incubating on ice for 1 hour; centrifugation for 1 hour – 14,000 rpm at 0°C.
17. Decanting the supernatant and adding 1 ml 90% acetone. Vortexing for 30 seconds and centrifuging for 10 minutes at 14,000 rpm. Decanting the supernatant.
18. Drying with a SpeedVac for 15 seconds.
19. Protein gel electrophoresis (SDS-PAGE - 10% gel, 20 cm long) and silver staining.

[0045] A gel image of high salt eluate from immobilized RNA Polymerase II from Saccharomyces cerevisiae and another eluate containing the core complex is shown in Figure 11. The carboxy-terminus of the chromosomal copy of the gene encoding Rpb1 is fused in-frame with a TAP-tag. 9 liters Saccharomyces cerevisiae culture was grown in YPD medium to OD_{600} = 0.8. The protocol used for obtaining the high salt eluate shown in the right lane is the same as in

Figure 10 except for the high salt elution that was performed with 0.3 M KCl in the washing buffer. After separating the interacting proteins by high salt elution and thus completing the method of the invention, the immobilized RNA Polymerase II complexes were released by treatment with 5mM EGTA and 0.5% SDS in WB. Half of the eluate was loaded in the left lane and subjected to PAGE-SDS in order to demonstrate the ratio between the amounts of the interacting proteins and the immobilized core complex. All gel bands, including the "blank" ones, were cut for analysis and the proteins that were identified with high confidence are indicated. Note the complete separation of the interacting proteins from the immobilized core complex by increasing the ionic strength (except TFIIF, which binds strongly to the core enzyme and was detected in both fractions). Different intensities of the bands in the two gel lanes illustrate the need to separate the substoichiometrically interacting proteins from the immobilized core complex. Note that a double gel well was used to load only a half of the SDS/EGTA eluate in order to avoid overloading and smearing of the gel. Legend: 271 – Spt6, 272 – Ydl145 (Cop1), 273 – Ctr9 + Spt5, 274 – Tfg1 + Spt16 + Lte1, 275 – Tfg1 + Spt5 + Spt16 + Fcp1, 276 – Set2, 277 – Fcp1 + Leo1, 278 – Pob3, 279 – Paf1, 280 – Paf1 + Rtt103, 281 – Tfg2 + Iws1, 282 – Tfg2, 283 – Cdc73 + Ceg1, 284 – Dst1, 285 – Rtt103, 286 – Tfg3, 287 – Ess1, 288 – Rpb1, 289 – Rpb2, 290 – Tfg1, 291 – Tfg2, 292 – Rpb3, 293 – Tfg3, 294 – Rpb4, 295 – Rpb5, 296 – Rpb6, 297 – Rpb7, 298 – Rpb8. Again, the separation of the immobilized complex from the affinity matrix is not necessary for practicing the invention, it is outside the scope of the invention and, when practicing the invention, the greatest care must be taken not to separate the immobilized fusion protein from the affinity matrix during the separation of the interacting proteins from the immobilized fusion protein.

[0046] The contaminants that were determined by performing the method of the invention with two immobilizations on two different affinity matrices, i.e. IgG-sepharose and calmodulin beads, and with cellular lysate from organism devoid of affinity tagged protein, are: Tef2, Fks1, Fas2, Stm1, Rpl4, Tif34, Rpl5, Rps3, Rps1b, Rpl2, Prt1, Rpl8, Tef2, Rps2, Rps5, Rpl10, Ded1, Rpl9, Hsp60, Rps8, Gcd11, Rps20, Rpl20, Pab1, Rps17, Ssb1, Ssa2, Sse1, Rpl27, Rps14, Rps24, Mdn1, Tra1, Dyn1, Myo2, Ira2.

[0047] Although the usage of dual affinity tag results in low background from the residual contaminant proteins, the number of the detected interacting proteins is lower than the number of

interacting proteins that are detected when a fusion protein with one tag (SpA-tag) is used. It is clear that the basic method of the invention (i.e. the protein of interest is fused to one affinity tag and only one immobilization of the protein of interest is performed before the separation of the interacting proteins) is faster, cheaper and results in identifying more interacting proteins. Compare the interacting proteins in Figure 3, 4, 6, 7 and 8 with Figure 9 and 10. Therefore, when the protein of interest is present in the cell in approximately 5,000 - 10,000 copies or higher, it is best if the basic method of the invention is carried out with a fusion protein containing one affinity tag. A good explanation for the loss of the interacting proteins when two sequential immobilization are performed is given by R. Aebersold and M. Mann, Nature, March 2003, Vol. 422. The background from the contaminant proteins creates problems mostly for the identification and analysis of the interacting proteins by mass spectrometry. If they are analyzed by other methods, e.g. immunoblotting, the background can be ignored. Also, the background from the contaminant proteins is negligible when compared to the background from the affinity tagged protein (or other subunits of the stable protein complex) that is encountered when a classical purification via affinity tag is performed.

[0048] The following proteins were identified by performing the method of the invention with several affinity tagged subunits of RNA Polymerase II from Saccharomyces cerevisiae. Proteins separated by a slash are subunits of stable protein complexes.
(a) transcription factors that are known to bind to RNA Polymerase II and proteins associated with them: Tfg1/Tfg2/Tfg3 – TFIIF, Spt5, TFIIS, Spt6/Iws1, Fcp1, Cet1/ Ceg1, TFIIB, Rgr1/Srb4/Med1/Med4 – mediator subunits, Ppn1 - interact with TFIIB, Nip1 - interact with TFIIB, Rtf1/Paf1/Ctr9/Leo1/Cdc73 – PAF complex, Spt16/Pob3, Taf7/Taf12 – TFIID, Ess1, Rad3/Tfb2 – TFIIH, Tfa2 (TFIIE).
(b) Proteins that are known to be involved in mRNA transcription or mRNA processing but are not known to associate directly or indirectly with the core enzyme: Npl3 -DNA binding; involved in mRNA transport, Hrp1 - mRNA cleavage/polyadenylation factor, Prp43 - mRNA splicing factor, Syf1 - mRNA splicing factor, Smb1 - mRNA splicing factor, Puf2 - mRNA binding protein, Gcd14 - RNA processing/modification factor, Fip1 - mRNA polyadenylation factor, Clp1 - cleavage-polyadenylation factor IA subunit, Ylr419 - RNA helicase; involved in RNA splicing, Rgm1 - transcription factor, Rgt1 - transcriptional activator, Cdc39 – transcription factor, Ygl244 - transcription cofactor activity, Mbf1 - bridges Gcn4 and Spt15, Stb5 - transcriptional activator,

(c) Proteins that interact with DNA/chromatin and/or mRNA and/or NTPs: Rap1- interacts with Rif1, DNA-binding protein, Rif2 - interacts with Rap1 and with Rif1, Ddr48 - DNA damage inducible protein, Ddc1 - involved in DNA damage checkpoint, Prp22 - helicase-like protein, Prp18 - helicase-like protein, Cdc6 - GTPase/ATPase; cell cycle control, Ygr054 - part of Swi/Snf and Npl3 complexes, Sdc25 - GDP/GTP exchange factor for Ras, Glo3 - GTPase activator,
(d) other proteins: Sap4 - associates with the SIT4 phosphatase, Sap185 - SIT4 associated protein, Ptc2 - protein phosphatase type 2C, Ubp8 - putative deubiquitinating enzyme, Kin4 - nuclear protein kinase, Hbt1 - Hub1 target, Yhr121 - Stb5 associated, Sup35 - interacts with Mip6 and Nab3, Sip1 - Snf1 protein kinase substrate, Rtt103 – binds to CTD, Mks1, Ydl145 (Cop1). More proteins are indicated in the Examples.
The nucleotide sequences of the genes and amino acid sequences of the proteins can be found by submitting the name of the protein at: Saccharomyces Genome Database at http://www.yeastgenome.org/ or, at National Center for Biotechnology Information web site at http://www.ncbi.nlm.nih.gov:80/mapview/map_search.cgi?taxid=4932 .

[0049] Note that by using only one method, i.e. the method of the invention, it is possible to identify directly nearly all proteins that are known to interact with RNA Polymerase II and that were identified during the last 20 years by using various methods (including genetic based and library based methods).

[0050] Note that among the detected proteins are two well-known enzymes that modify covalently the carboxyl-terminal domain (CTD) of RNA Polymerase II: (a) Fcp1 – a TFIIF interacting phosphatase that recycles RNA Polymerase II, (b) Ess1 – a prolyl isomerase of the CTD. Since post-translational modifications play a major role in modulating the protein function, the identification of the modifying enzyme for a protein of interest is an important application of the invention. In addition, the invention can be used to identify the substrate for an enzyme of interest.

[0051] The presence of many (more than 30) well-established transcription factors in the high salt eluates obtained from immobilized RNA Polymerase II is enough for validation of the invention, i.e. the proteins isolated by the method of the invention are true interacting proteins. One way to validate interacting proteins is to perform the method of the invention with a fusion protein

containing the putative interacting protein. The presence of the first protein of interest among the interacting proteins increases the confidence that the protein-protein interaction is physiologically relevant. For example, by performing the method of the invention with tagged subunits of RNA Polymerase II, the three subunits of the transcription factor TFIIF (Tfg1, Tfg2 and Tfg3) are detected in the high salt eluate, and, on the other hand, by performing the method of the invention with tagged Tfg1, the subunits of the RNA Polymerase II core enzyme are detected (Figures 7 and 8).

[0052] Alternatively, the validation can be performed by other methods for detection of protein interactions as described by Phizicky E. and Fields S. in Microbiological Reviews, Mar. 1995, Vol. 95, No. 1, or as described by Phizicky E, Bastiaens PI, Zhu H, Snyder M, Fields S., Nature 2003. For example, purification via an affinity tag fused to the putative interactor can be performed. Rtt103 is among the interacting proteins detected by the method of the invention and, in order to validate it as an interactor, a tandem affinity purification was performed with a strain expressing TAP-tagged Rtt103. As shown in Figure 21, two subunits of the RNA Polymerase II core complex (Rpb1 and Rpb2) co-purified with Rtt103. Legend: 470 - Rpb1, 471 - Rpb2, 472 – Rat1, 473 - Rtt103, 474 - Rai1.

[0053] The method of the invention will be especially useful for detecting protein interactions in organisms where genetic approaches are impossible or very hard (e.g. Homo sapiens). The invention will help to elucidate the complete set of interactions that involve proteins having associations with human diseases and will reveal new targets for therapeutic intervention.

[0054] The eluate obtained by the method of the invention can be analyzed for the presence of substances that can be drugs or pre-drugs. In addition, a drug library can be applied to the living cells and/or the cellular lysate and the eluate can be analyzed for the presence of one or more members of the library. Also, after the elution of the interacting proteins, the immobilized protein or protein complex can be used for screening libraries of chemicals for a particular chemical that binds to it. Methods for drug discovery are described by Rossi D. and Sinz M., Marcel Dekker 2001 "Mass Spectrometry in Drug Discovery".

[0055] Moreover, the invention can be used as a novel method for drug design and drug discovery. A chemical can be identified as a drug or a pre-drug by its capability to affect the formation and/or disintegration of a particular protein complex that is associated with a disease. Such a chemical should have two characteristic features: (a) it should affect selectively a particular protein interaction (or group of interactions involving a protein that is associated with a disease), unlike increasing the ionic strength which affects transient protein interactions nonselectively, (b) it should affect the particular protein interaction at a concentration below 10 mM, i.e. a concentration that does not change significantly the ionic strength. After the chemical or a biomolecule is identified as a pre-drug, it can be further modified and the experiment can be repeated with lower concentration of the said chemical or biomolecule. The goal is to obtain such a chemical that affects selectively the formation and/or disintegration of the protein complex at micromolar concentration, i.e. concentration that is appropriate for treatment of human subjects.

[0056] The effect of the drug or pre-drug can be determined by monitoring the presence and/or the concentration of the interacting proteins in the eluate. Preferably, after building a proteomic map around the protein of interest (i.e. identifying unambiguously the interacting proteins), other methods, that are more sensitive than mass spectrometry, should be used. For example, antibodies can be raised against a particular protein and it can be detected by immunoassays. Also, the interacting proteins can be tagged and detected by antibodies against the tag.

[0057] The effect of the pre-drug on a particular protein interaction can be tested in several ways:
(i) The pre-drug can be added to the growing cells, the lysis buffer and the washing buffer. In this case, the elution is performed with washing buffer without a pre-drug. Alternatively, the elution is performed with washing buffer with increased ionic strength but without a pre-drug.
(ii) When the chemical is applied to the immobilized protein complex after the removal of unbound substances, it can be used as a single elution agent or together with one or more nonspecific elution agents, i.e. the elution can be done with the chemical in the washing buffer or with the chemical in the elution buffer.
(iii) The high salt eluate, optionally dialyzed against washing buffer or other physiological buffer, can be recombined with the immobilized fusion protein and the pre-drug and incubated for 10-60 minutes or more, depending on the concentration of the proteins and the equilibrium binding constant of the protein interaction. The restoration of the protein complex can be determined by

monitoring the amount of one or more interacting proteins in the eluate and/or on the affinity matrix (i.e. after the incubation and the removal of unbound proteins, the elution is performed with high salt buffer).

[0058] A method for drug design and drug discovery that is based on the finding that the nature of transient protein-protein interactions is dominantly electrostatic contains the following steps:
(i) Selecting a protein or a protein complex associated with a disease and obtaining a cell line or organism with an affinity tagged gene(s). According to our finding, a disease that is caused by a mutation that changes the electrostatic properties of a protein is due to an aberration in a transient protein interaction. The mutation can change the electrostatic properties of the protein by replacing Lysine and/or Arginine and/or Aspartate and/or Glutamate and/or Histidine with an uncharged or oppositely charged amino acid. In addition, a mutation can change the electrostatic properties of the protein by replacing other amino acids with Lysine and/or Arginine and/or Aspartate and/or Glutamate and/or Histidine. See "Appendix – List of diseases" and note how many severe diseases are caused by such mutations.
(ii) Performing the method of the invention and/or other affinity purifications in order to sort out which group of proteins are associated in stable protein complexes (i.e. associated mainly by hydrophobic forces) and which ones are associated in transient complexes (i.e. associated mainly by electrostatic forces). Building a proteomic map (i.e. map of protein-protein interactions) around the protein that is associated with the disease.
(iii) Performing the method of the invention under different conditions (functional proteomics) and/or with cell lines originating from organism(s) with a disorder. Determining which protein interaction is associated with the specific disease and determining the structure and the composition of the interface (the part of a protein surface that is interacting with the other protein).
(iiii) Designing and/or synthesizing a chemical or a group of chemicals capable of binding to the mutant protein and restoring, at least partially, its electrostatic properties. Alternatively, the drug may be designed to bind to the native protein interactor and modify its electrostatic properties in such a way that it can interact with the mutant protein. An already existing drug or chemical can be used in the beginning and later can be modified. The drug should consist of two parts – (I) Part that binds selectively to the protein and ensures the specificity of the drug. Methods for designing and/or synthesizing chemicals that bind selectively to particular protein structure (most often by

shape complementarity) are described in the literature and known to persons skilled in the art - Pong S., Biopharmaceutical Drug Design and Development, Blackwell Publishing 1999, (II) Part that is electrostatically charged and restores, at least partially, the electrostatic properties of the protein. For example, if a mutation is replacing a positively charged amino acid (Lysine or Arginine) with uncharged one, the drug should contain a positively charged group and after the binding of the drug to the protein, such group should be located at the same place as the mutated amino acid or close to it. Designing a drug in such a way that it binds selectively to a mutant protein and brings in an electrostatic charge that modifies the electrostatic properties of the mutant protein is a novel approach for drug design.
(iiiii) – Testing the capability of the chemical (pre-drug) to affect the proper formation and/or disintegration of the protein complex that is associated with a disease.

[0059] Figure 12. Comparison between the method of the invention and other methods for detecting protein-protein interactions that utilize affinity tagged proteins. Legend: 301 - the left panel illustrates GST pull down, 302 - the middle panel illustrates purification via an affinity tag, 303 - the right panel illustrates the method of the invention. 310 – exogenous protein (e.g. Escherichia coli), 311 – endogenous protein, 312 – affinity tagged protein, 313 – the affinity tagged protein and the interacting proteins are immobilized on solid phase, 314 – unbound proteins and/or other substances (solid black) are removed, 315 – isolating the interacting proteins – only in the case of the invention they are separated from the affinity tagged protein (the eluate does not contain affinity tagged protein), 316 – eluate obtained by GST pull down; 317 – eluate obtained by purification via an affinity tag; 18 - eluate obtained by the method of the invention.
GST pull down – a single fusion protein is expressed in an exogenous organism (most often E. coli) and is used to prepare the affinity column. GST pull down can be performed by separating the fusion protein from the affinity matrix or by separating the interacting proteins from the immobilized protein of interest.
Purification via an affinity tag – the fusion protein and the interacting proteins are isolated together. Only in rare cases can transient complexes be detected.
The invention – the substoichiometrically interacting proteins are separated from the immobilized fusion protein and the dynamic range problem and/or other problems are eliminated.

[0060] Figure 13 illustrates the necessity to separate substoichiometrically interacting proteins from the high abundance fusion protein. The bulk of the fusion protein is immobilized via an affinity tag whereas only the fraction of interacting proteins that has been bound to the fusion protein in vivo is immobilized. 330 – transient complexes are immobilized on an affinity matrix via an affinity tag, 331 – purification via an affinity tag, 332 – purification by the method of the invention, 317 – eluate obtained by purification via an affinity tag, 18 - eluate obtained by the method of the invention. Note the amount of the tagged protein that is not associated with interacting proteins. In the illustration the ratio between the fusion protein and each interacting protein is 4:1 but in reality, the ratio can vary between 10:1 and more than 100:1. Moreover, when the fusion protein is a subunit of a stable protein complex, the problems arising from the different stoichiometries are even more complicated

[0061] Figure 14 illustrates the formation and disintegration of a transient protein complex between two stable (permanent) protein complexes (first and second vertical arrows respectively). Legend: 351 – stable protein complex with 4 subunits, 352 – stable protein complex with 6 subunits, 353 - transient protein complex between the two stable complexes, 354 - subunit of a stable complex that binds indirectly to another stable complex, 355 - formation of a transient protein complex, 356 - disintegration of the complex.

[0062] Figure 15. A model for transient protein-protein interactions according to which formation and disintegration of transient protein complexes is due to alternation of electrostatic attraction and electrostatic repulsion between the electrostatically charged amino acids.
Legend: 370 – first interacting protein, 371 – second interacting protein, 372 – positively charged amino acids, 373 – negatively charged amino acids, 374 – electrostatic attraction, 375 – two interacting proteins form a complex, 376 - change of position or orientation of the two interacting proteins relative to one another, 377 – electrostatic repulsion, 378 – transient complex disintegrates.

[0063] Figure 16. When a mutation changes the electrostatic properties of the protein, the transient interaction cannot occur properly. Legend for the upper panel: 380 – mutation leading to disappearance of positively charged amino acid, 381 – electrostatic attraction is prevented. Legend for the lower panel: 382 - mutation leading to disappearance of different positively

charged amino acid, 383 - electrostatic repulsion is prevented and the interacting proteins remain stuck together.

[0064] Figure 17. Method for drug design and schematic structure of a chemical or a biomolecule that binds to a mutant protein and restores (at least partially) its electrostatic properties. As a result, the protein is capable of forming a protein complex. Legend: 390 – drug treatment, 391 – a drug, 392 – an exploded view of the drug, 393 - part that mimics and/or restores the original charge, 394 - linker (optional), 395 - part that binds selectively to the mutant protein.

[0065] When the disease is caused by a mutation that converts a permanent protein-protein association into a weak interaction, the drug should stabilize the protein-protein association. Drugs that are designed or discovered according to the invention, can be defined as chemicals that restore the equilibrium binding constant of a mutant protein complex back to that of the respective native one.

[0066] If a human disease has an analog in another organism (e.g. mouse or rat), a heterologous nucleic acid can be administered into the organism so that it expresses an affinity tagged protein of interest. In this case, the method of the invention can be performed with cellular lysate or other biological fluids from the sacrificed organism.

Examples of protocols for carrying out the invention with different affinity tags. The scale of the experiments is suitable for detection by mass spectrometry of proteins interacting with a protein of interest that is present in approximately 10,000 copies per cell or more. The examples are illustrative but not limiting the scope of the invention.

[0067] After introducing a heterologous nucleic acid, Saccharomyces cerevisiae strain expresses a fusion protein containing a GST-tag. The following protocol can be used:
Growing 12 liters Saccharomyces cerevisiae culture in YPD medium to $OD_{600} = 0.8$.

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations are performed on ice or at 0°C to 5°C.

2. Suspending the cells in 100 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 100mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine.
5. Centrifugation at 16,000 rpm for 30 minutes and transferring the supernatant to a clean tube.
6. Adding 500 microliters of glutathione beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the glutathione beads to two parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 50mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 4ml WB with 0.4M KCl – WB (0.4M)
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.4M) and combining the eluates.
12. Analyzing the interacting proteins and/or other biomolecules in the eluate. Optionally, proceed to 14.
13. Distributing the eluate to eppendorf tubes and adding 1/5 volume of 100% TCA.
14. Incubating on ice for 30 minutes; centrifugation for 30 minutes – 14,000 Rpm at 0°C to 5°C.
15. Decanting the supernatant and adding 1 ml 90% acetone. Vortexing for 30 seconds and centrifuging for 10 minutes at 14,000 rpm. Decanting the supernatant.
16. Drying with SpeedVac for 15 seconds.
17. Protein gel electrophoresis (SDS-PAGE – 10% gel, 20 cm long) and silver staining.

[0068] When the interacting proteins are isolated in low amounts and an additional affinity purification step is needed before the separation from the immobilized fusion protein, a variety of combinations of two affinity tags can be used. For example, the combination can be a GST-tag

and a SpA-tag. The following protocol can be used in several cases: (a) One subunit of protein complex is fused to a GST-tag and another subunit is fused to a SpA-tag, (b) The protein or subunit of the protein complex contains a GST-tag at the N-terminus and a SpA-tag at the C-terminus, (c) Both tags can be located at the same end.

Growing 18 liters Saccharomyces cerevisiae culture in YPD medium to $OD_{600} = 0.8$.

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations are performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine.
5. Centrifugation at 16,000 rpm for 30 minutes and transferring the supernatant to a clean tube.
6. Adding 800 microliters of glutathione beads and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the glutathione beads to four parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 3ml WB containing 10 mM reduced glutathione.
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB containing 10 mM reduced glutathione and combining the eluates in a 5 ml tube.
12. Adding 600 microliters of IgG beads to the eluate and rotating the tube for 1 hour.
13. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
14. Transferring and distributing the IgG beads to two parallel 10ml chromatography columns. Washing each column with 10ml WB by applying 200 microliter aliquots.

15. Transferring the beads to a 5ml tube and adding 3ml WB containing 0.5M KCl - WB(0.5M). Rotating the tube for 10 minutes.
16. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.5M) and combining the eluates.
17. Analyzing the eluate.

[0069] Another example is a combination of a SpG-tag and a SBP (Streptavidin binding peptide - MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP). The following protocol can be used:

Growing 12 liters Saccharomyces cerevisiae culture in YPD medium to $OD_{600} = 0.9$.

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations are performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes, centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine.
5. Centrifugation at 16,000 rpm for 20 minutes and transferring the supernatant to a clean tube.
6. Adding 600 microliters streptavidin beads and rotating the tube for 1.5 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the streptavidin beads to three parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 4ml WB containing 4mM biotin.
10. Rotating the tube for 10 minutes.

11. Transferring the beads and the buffer to 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB containing 4mM biotin and combining the eluates in a 5 ml tube.
12. Adding 600 microliters of IgG beads to the eluate and rotating the tube for 1 hour.
13. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
14. Transferring and distributing the IgG beads to three parallel 10ml chromatography columns. Washing each column with 10ml WB by applying 200 microliter aliquots.
15. Transferring the beads to a 5ml tube and adding 3.5 ml WB containing 0.5M KCl - WB(0.5M). Rotating the tube for 10 minutes.
16. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.5M) and combining the eluates.
17. Analyzing the eluate.

[0070] Another example is a combination of a MBP and a HA-tag (YPYDVPDYASL). The following protocol can be used:

Growing 12 liters Saccharomyces cerevisiae culture in YPD medium to $OD_{600} = 0.9$.

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations are performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes, centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2 minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine.
5. Centrifugation at 16,000 rpm for 20 minutes and transferring the supernatant to a clean tube.
6. Adding 800 microliters of anti-HA affinity matrix and rotating the tube for 2 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the anti-HA affinity matrix to four parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying

200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.

9. Transferring the beads to a 5 ml tube and adding 3 ml WB containing HA-tag at 1 mg/ml.
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB containing HA peptide at 1 mg/ml and combining the eluates in a 10 ml tube.
12. Adding 600 microliters of cross-linked amylose beads to the eluate and rotating the tube for 1 hour.
13. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
14. Transferring and distributing the beads to three parallel 10ml chromatography columns. Washing each column with 10ml WB by applying 200 microliter aliquots.
15. Transferring the beads to a 5ml tube and adding 3.5 ml WB containing 0.5M KCl - WB(0.5M). Rotating the tube for 10 minutes.
16. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.5M) and combining the eluates.
17. Analyzing the eluate.

[0071] Another example is a combination of a SBP (Streptavidin binding peptide-MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP) and a 6xHis tag.
The following protocol can be used:
Growing 18 liters Saccharomyces cerevisiae culture in YPD medium to $OD_{600} = 0.7$.

1. Collecting the cells by centrifugation at 4,000 rpm for 4 minutes and decanting the medium. All subsequent manipulations are performed on ice or at 0°C to 5°C.
2. Suspending the cells in 200 ml water and distributing into 50ml plastic tubes; centrifugation at 4,000 rpm for 4 minutes and decanting the water.
3. Freezing in liquid nitrogen for 1 minute.
4. Breaking the solid pellet and adding protease inhibitors and 2-3 pieces (around 8 cm^3) dry ice; grinding with a coffee grinder for 1 minute; transferring the powder to a beaker and, after allowing it to melt, adding the same volume 2x Lysis buffer. Stirring with a plastic rod for 2

minutes. 2x Lysis Buffer: 200mM KCl, 50mM HEPES pH 7.4, 20% glycerol, 2% DMSO, 0.2% TritonX100, 2mM PMSF, 2mM benzamidine.

5. Centrifugation at 16,000 rpm for 20 minutes and transferring the supernatant to a clean tube.
6. Adding 800 microliters streptavidin beads and rotating the tube for 1.5 hours.
7. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
8. Transferring and distributing the streptavidin beads to four parallel 10ml chromatography columns. Washing each column with 10ml washing buffer (WB) by applying 200 microliter aliquots. WB: 100mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100.
9. Transferring the beads to a 5ml tube and adding 4ml WB containing 4mM biotin.
10. Rotating the tube for 10 minutes.
11. Transferring the beads and the buffer to 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB containing 4mM biotin and combining the eluates in a 10 ml tube.
12. Adding 600 microliters of Ni^{2+}-NTA resin to the eluate and rotating the tube for 1 hour.
13. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
14. Transferring and distributing the Ni^{2+}-NTA resin to three parallel 10ml chromatography columns. Washing each column with 10ml WB by applying 200 microliter aliquots.
15. Transferring the Ni^{2+}-NTA resin to a 5ml tube and adding 3.5 ml WB containing 0.4M KCl - WB(0.4M). Rotating the tube for 10 minutes.
16. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.4M) and combining the eluates.
17. Analyzing the eluate.

[0072] When TAP-tag is used and the first immobilization is on calmodulin beads, the lysis buffer and washing buffer should contain 2mM Ca++. After removing the unbound substances and releasing the fusion protein by 2mM EGTA treatment, the fusion protein is immobilized on IgG beads and the residual unbound substances are removed. The separation of the interacting proteins and/or other substances from the immobilized fusion protein is performed with WB containing 0.6 M KCl.

Examples of carrying out the invention with a protein of interest devoid of an affinity tag:

[0073] A specific antibody can be raised against the protein of interest and subsequently immobilized on a solid support. Cellular lysate or a fraction of cellular lysate from an organism lacking an affinity tagged protein is mixed with antibody coated solid support and after removing the unbound substances, biomolecules that are bound to the protein of interest are separated from it by increasing the salt concentration to 0.3M KCl. The following protocol can be used (all manipulations are performed at 0°C to 5°C):
The cells are collected and frozen in liquid nitrogen for 1 minute. After melting and adding lysis buffer, the lysate is homogenized by 10-20 strokes.

1. Cellular lysate from 10^{10} mammalian cells in 10ml Lysis Buffer (50mM KCl, 10mM HEPES pH 7.4, 10% glycerol, 1% DMSO, 0.1% TritonX100, 1mM PMSF, 1mM benzamidine) is centrifuged at 15,000 rpm for 30 minutes and the supernatant is transferred to a clean tube.
2. Adding 200 microliters of antibody coated beads and rotating the tube for 1.5 hours.
3. Centrifugation at 2,000 rpm for 3 minutes and decanting the supernatant.
4. Transferring the beads to a 10ml chromatography column. Washing the column with 10ml washing buffer (WB) by applying 200 microliter aliquots (WB: 50mM KCl, 10mM HEPES pH 7.4, 7% glycerol, 1% DMSO, 0.1% TritonX100).
5. Transferring the beads to a 5ml tube and adding 3ml WB containing 0.3M KCl – WB(0.3M)
6. Rotating the tube for 10 minutes.
7. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.3M) and combining the eluates.
8. Analyzing the eluate.

[0074] The previous protocol can be modified in order to avoid the immobilization of the antibody. An IgG type antibody can be raised against the protein of interest and, after binding to the corresponding protein of interest, the antibody can be immobilized on Protein A - Sepharose. After removal of the unbound substances, biomolecules that are bound to the protein of interest are separated from it by increasing the salt concentration to 0.35M KCl. The following protocol can be used (all manipulations are performed at 0°C to 5°C):

The cells are collected and frozen in liquid nitrogen for 1 minute. After melting and adding lysis buffer, the lysate is homogenized by 10-20 strokes.

1. Cellular lysate from $5x10^9$ mammalian cells in 10ml Lysis Buffer (50mM KCl, 10mM HEPES pH 7.4, 10% glycerol, 1% DMSO, 0.1% TritonX100, 1mM PMSF, 1mM benzamidine, 0.5 mM $MgCl_2$, 0.1 mM DTT) is centrifuged at 15,000 rpm for 20 minutes and the supernatant is transferred to a clean tube.
2. Adding the antibody and incubating for 1.5 hours.
3. Adding 300 microliters of Protein A-sepharose beads and rotating the tube for 1hour.
4. Centrifugation at 2,000 rpm for 3 minutes and decanting the supernatant.
5. Transferring the beads to a 10ml chromatography column. Washing the column with 10ml washing buffer (WB) by applying 200 microliter aliquots (WB - 50mM KCl, 10mM HEPES pH 7.4, 7% glycerol, 1% DMSO, 0.1% TritonX100, 0.5 mM $MgCl_2$, 0.1 mM DTT).
6. Transferring the beads to a 5ml tube and adding 3ml WB containing 0.35M KCl – WB(0.35M)
7. Rotating the tube for 15 minutes.
8. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB (0.35M) and combining the eluates.
9. Analyzing the eluate.

[0075] Another enhancement consists of using biotinylated antibodies and utilizing their strong binding to streptavidin. The antibody can be biotinylated and added to the cellular lysate and, after 1 hour incubation, the lysate can be mixed with streptavidin coated sepharose beads. The following protocol can be used (all manipulations are performed at 0°C to 5°C):

1. The cells are collected and frozen in liquid nitrogen for 1 minute. After melting and adding lysis buffer, the lysate is homogenized by 10-20 strokes.
2. Cellular lysate from 10^{10} mammalian cells in 10ml Lysis Buffer (50mM KCl, 10mM HEPES pH 7.4, 10% glycerol, 1% DMSO, 0.1% TritonX100, 1mM PMSF, 1mM benzamidine, 0.5 mM $MgCl_2$, 0.1 mM DTT) is centrifuged at 16,000 rpm for 30 minutes and the supernatant is transferred to a clean tube.
3. Adding biotinylated antibody and rotating the tube for 1.5 hours.
4. Adding 200 microliters of streptavidin-sepharose beads and rotating the tube for 1 hour.

5. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
6. Transferring the beads to 10ml chromatography column. Washing the column with 10ml washing buffer (WB) by applying 200 microliter aliquots (WB - 50mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100).
7. Transferring the beads to a 5ml tube and adding 4ml WB containing 0.3M KCl – WB(0.3M)
8. Rotating the tube for 10 minutes.
9. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.3M) and combining the eluates.
10. Analyzing the eluate.

[0076] Proteins that interact with glycoproteins can be isolated by the method of the invention. Cellular lysate or a fraction of cellular lysate or other biological fluid from an organism or cell culture without a genetically modified gene (i.e. devoid of an affinity tag) is mixed with lectin-coated beads in order to immobilize the glycoproteins. After removing the unbound substances, biomolecules that interact with glycoproteins are separated from them by increasing the salt concentration to 0.4 - 0.5 M KCl. In this case, what is isolated is not a group of proteins that interact with a particular glycoprotein, but proteins or other substances that interact with any glycoprotein that is immobilized on the lectin matrix. Since glycolipids and polysaccharides also bind to lectin, it is recommended to perform one purification step before the method of the invention in order to obtain a fraction enriched in glycoproteins. The following protocol can be used (all manipulations are performed at 0°C to 5°C):

1. The cells are collected and frozen in liquid nitrogen for 1 minute. After melting and adding lysis buffer, the lysate is homogenized by 10-20 strokes.
2. Cellular lysate from 10^{10} mammalian cells in 10ml Lysis Buffer (50mM KCl, 10mM HEPES pH 7.4, 10% glycerol, 1% DMSO, 0.1% TritonX100, 1mM PMSF, 1mM benzamidine, 1mM $MnCl_2$, 1mM $CaCl_2$) is centrifuged at 16,000 rpm for 30 minutes and the supernatant is transferred to a clean tube. (Optionally, the lysate can be fractionated in order to obtain a fraction that is enriched in glycoproteins or a particular glycoprotein).
1. Adding 200 microliters of lectin coated beads and rotating the tube for 1 hour.
2. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.

3. Transferring the lectin beads to a 10ml chromatography column. Washing the column with 10ml washing buffer (WB) by applying 200 microliter aliquots (WB - 50mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 1% DMSO, 0.1% TritonX100, 1mM $MnCl_2$, 1mM $CaCl_2$).
4. Transferring the beads to a 5ml tube and adding 3.5ml WB containing 0.4M KCl – WB(0.4M)
5. Rotating the tube for 10 minutes.
6. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB(0.4M) and combining the eluates.
7. Analyzing the eluate.

Examples of using the invention for drug discovery

[0077] The invention can be used for monitoring the effect of a pre-drug on a protein-protein interaction that is associated with a disease. The protein can be fused to a SpA-tag but if the cellular lysate or other biological fluid contains a high level of immunoglobulins or other proteins that bind to SpA-tag or IgG, a different affinity tag should be used (e.g. MBP or SBP or GST). In this example, the protein does not contain a pathogenic mutation. It is best if a complete proteomic map has been built around the protein of interest before the experiment, so that the presence of interacting proteins in the eluate can be determined by immunoassays or other sensitive methods.

Cellular lysate or a fraction of cellular lysate or other biological fluid from human cells containing a modified gene that encodes a fusion protein containing the SpA-tag is mixed with IgG-beads. After removing the unbound substances, the immobilized complex is treated with washing buffer containing the pre-drug in concentration between 0 and 1mM. The concentration of the pre-drug can vary. The effect of the pre-drug is determined by monitoring the presence of interacting proteins in the eluate. The following protocol can be used (all manipulations are performed at 0°C – 5°C):

1. The cells are collected and frozen in liquid nitrogen for 1 minute. After melting and adding lysis buffer, the lysate is homogenized by 10-20 strokes. Cellular lysate from $5x10^9$ mammalian cells in 10ml Lysis Buffer (50mM KCl, 10mM HEPES pH 7.4, 10% glycerol, 1%

DMSO, 0.1% TritonX100, 0.2mM PMSF, 0.2mM benzamidine,) is centrifuged at 16,000 rpm for 20 minutes and the supernatant is transferred to a clean tube.

2. Adding 100 microliters of IgG beads and rotating the tube for 1.5 hours.
3. Centrifugation at 3,000 rpm for 2 minutes and decanting the supernatant.
4. Transferring the beads to a 10 ml chromatography column. Washing the column with 10ml washing buffer (WB) by applying 200 microliter aliquots (WB - 50mM KCl, 10mM HEPES pH 7.4, 5% glycerol, 0.2% DMSO, 0.1% TritonX100, 0.2mM PMSF, 0.2mM benzamidine).
5. Transferring the beads to a 2ml tube and adding 0.5 ml WB containing 1mM pre-drug.
6. Rotating the tube for 30 minutes.
7. Transferring the beads and the buffer to a 10ml chromatography column and collecting the eluate by gravity flow. Washing the beads with two column volumes of WB containing1mM pre-drug and combining the eluates.
8. Analyzing the eluate for the presence or absence of interacting proteins or other biomolecules in order to determine the effect of the pre-drug on the complex.

The above method can be modified in several ways. (I) The fusion protein contains a pathogenic mutation and the pre-drug is present in all solutions: growing medium, lysis buffer and washing buffer. In steps 5, 6 and 7 the elution is performed by WB without a pre-drug in order to check whether the absence of the pre-drug leads to dissociation of the complex. (II) Another variant is to perform the treatment in steps 5, 6 and 7 with WB containing 0.3M KCl and a pre-drug. Presumably, the concentration of the drug should be lower when it is applied to living cells (i.e. in vivo) and higher when it is applied to the buffers (i.e. in vitro).

[0078] If the pre-drug affects the complex at 1 mM concentration, the experiment can be repeated with lower concentrations of the pre-drug, e.g. 100mcM or 10 mcM or 1mcM, in order to determine the lowest concentration that affects the complex. After conducting several experiments, the pre-drug can be modified and experiments can be repeated in order to determine whether the modification increases the specificity of the pre-drug, i.e. whether the pre-drug affects the complex at lower concentrations.

[0079] Preferably, before conducting the drug treatment experiments, a proteomic map should be built around the protein of interest so that it is known which protein-protein interaction is

associated with a particular disease. In this case, both proteins can be tagged and the interacting protein can be detected by immunoassays or other assays in order to reduce the scale of the experiment. Columns containing as low as 10 mcl or 20 mcl affinity matrix can be used (Formosa T., Methods Enzymol. 1991). Alternatively, if both proteins are present in a very low copy number in the cell, they can be over-expressed as fusion proteins in an exogenous or endogenous organism and the experiment can be performed with purified proteins. Preferably, the proteins can be over-expressed in their endogenous organism and the regular method of the invention should be carried out. Again, the over-expression should not lead to formation of inclusion bodies.

[0080] Two types of control experiments can be performed in order to validate the effect of the pre-drug: (I) step 5, 6 and 7 are performed without a pre-drug, (II) testing the effect of a pre-drug on one or more biologically unrelated protein complexes.

[0081] Figure 18 illustrates the selective effect of a drug on a particular protein-protein interaction. Legend: 430 – left panel: the immobilized protein of interest is treated with nonspecific elution agent and all interacting proteins are separated, 431 – middle panel: the immobilized protein of interest is treated with specific elution agent which causes separation of only one interacting protein complex, 432 – right panel: the unrelated protein complex is treated with the same specific elution agent but there is no separation of the interacting proteins, 438 – unrelated protein, 439 – protein interacting with the unrelated protein, 440 – another protein interacting with the unrelated protein, 440 – elution with nonselective elution agent (e.g. increased salt concentration), 441 - elution with selective agent (e.g. drug treatment), 18 – eluate obtained by the method of the invention, i.e. proteins eluted by treatment with a nonselective elution agent, 445 – proteins eluted by treatment with a selective agent; 446 – the unrelated protein complex is not affected by the treatment with a selective agent and the interacting proteins are not eluted.

Additional considerations about the invention

[0082] It is good if the affinity matrix is in the form of beads and can be packed into a column so that the removal of the unbound proteins and other substances can be achieved by washing the column. Elution by ionic strength equivalent to 0.35M KCl for 20 minutes is a good starting point

for studying an uncharacterized protein complex. Preferably, the affinity tag should bind the ligand-coated matrix by predominantly hydrophobic forces and the elution of the interacting proteins should be carried out by increasing the ionic strength and/or the dielectric constant of the medium. The elution should be performed under conditions that do not weaken the binding between the affinity tag and the ligand or weaken them to a lesser extent than the binding between the fusion protein and the interacting proteins and/or other substances.
Another way to immobilize the protein of interest is by metal chelate affinity (e.g. poly-His tag) and to elute the interacting proteins by increasing the ionic strength of the medium.
In addition, the protein of interest can be immobilized by predominantly electrostatic forces (e.g. by Flag-tag) and the elution can be performed with an agent that affects the strength of hydrophobic bonds (e.g. ethylene glycol).

[0083] The scale of the experiment (number of the cells and the amount of total protein) depends on several factors: (a) the copy numbers of the protein of interest and the interacting proteins, (b) the binding equilibrium constants of particular protein interactions, (c) the sensitivity of the method that is used to identify the interacting proteins. By performing the method of the invention with different tagged subunits and with approximately 10^{11} cells, nearly all reported interacting proteins for RNA Polymerase II (approximately 30,000 copies per cell) are detected by mass spectrometry. The copy number of the protein of interest can be determined by Western blotting and the scale of the experiment can be recalculated.

[0084] It is good to use mechanical forces for preparation of the cellular lysate, e.g. grinding the frozen cells or glass beads beating or brief sonication or homogenization. Centrifugation can be ultracentrifugation (above 15,000 rpm). Temperature should be kept between 0°C and 10°C and the pH should be optimally close to the physiological one.

[0085] It is a variant to add in the elution buffer substances that facilitate the separation of the interacting proteins from the protein of interest (e.g. 0.5M urea or 0.5M guanidinium chloride or SDS below 0.1% or TritonX100 below 1%) but do not disrupt the bond between the affinity tag and the affinity matrix.

[0086] Different analytical methods can be used to identify and analyze the proteins that are purified by the method of the invention, e.g. mass spectrometry, enzymatic assays, immunodetection. Immunodetection can be used to analyze the interacting proteins and/or to verify the identity of the interacting proteins that cannot be identified unambiguously by mass spectrometry. When the proteins are eluted with ionic strength equivalent to 0.3M KCl or 0.4M KCl, they are not denatured and the eluate can be analyzed by biochemical assays for the presence of a particular enzymatic activity. Optionally, the eluate can be dialyzed against physiological buffer before the assay.

[0087] After building a proteomic map around the protein of interest (i.e. identifying unambiguously the interacting proteins), more sensitive methods for detecting the interacting protein can be used. For example, antibodies can be raised against particular interacting protein and it can be detected by immunodetection. However, one of the main utilities of the invention is the direct identification (preferably by mass spectrometry) of novel proteins that interact with a protein of interest in vivo.

[0088] After separating the interacting proteins in the liquid phase, they can be further fractionated by SDS-PAGE (the length of the gel depends on the complexity of the eluate) and localized by silver staining or Coomassie staining. Another variant is to run a native gel and separate different permanent complexes from one another.

[0089] Elution can be performed by gradually increasing the concentration of the elution agent (e.g. salt gradient or pH gradient). Another variant is to fractionate further the proteins and/or other biomolecules in the eluate by chromatography (e.g. liquid chromatography) or other separating techniques. However, the fractionation may lead to unnecessary dilution and/or distribution of an interacting protein in several fractions. In this case, all the fractions that do not contain the affinity tagged protein, or contain only insignificant amount, can be recombined.

[0090] The elution agent might disrupt partially and nonselectively the binding between the affinity tag and the corresponding ligand (especially if some affinity tags are partially denatured during the preparation of the cellular lysate) and, as a result, a certain amount of the immobilized fusion protein can detach from the affinity matrix and co-elute with the interacting proteins. The

problem can be solved in several ways: (i) The affinity matrix can be pre-treated under the same conditions as the elution conditions in order to wash away any ligand that is not attached strongly enough to the solid support (e.g. agarose or sepharose), (ii) The highest concentration of the elution agent that does not cause separation of the fusion protein from the affinity matrix can be determined in a pilot experiment(s) using the same affinity tag fused to a standard (control) protein. Different fractions can be analyzed by immunodetection (with antibody against the affinity tag or against the standard protein) or other techniques in order to determine the presence or absence of the affinity tag and/or affinity tagged standard protein. The elution conditions that do not lead to separation of the affinity tagged standard protein from the affinity matrix should be used to perform the method of the invention with the protein of interest, (iii) The presence of an interacting protein in a fraction that is eluted before the first fraction that contains the fusion protein is sufficient to prove that the interacting protein is isolated and/or detected by separating it from the immobilized fusion protein. (iiii) If the fusion protein leaks together with the ligand (e.g. IgG), this is a problem with the affinity matrix. Figure 19 illustrates elution by a gradient and collecting different fractions and analyzing each one (by immunodetection for example) for the presence of the affinity tag or the affinity tagged protein: 401 - chromatography column; 402 – nine subsequent fractions, small amount of each one are blotted to a membrane and analyzed; 411, 412, 413, 414, 415, 416 - blots corresponding to fractions that do not contain affinity tagged protein, 417, 418, 419 – blots corresponding to fractions that contain increasing amounts of the affinity tagged protein.

[0091] Although the construction of affinity tags and expressing fusion proteins is not the essence of the present invention, several related issues deserve attention. When the organism is diploid, it is best to modify both chromosomal copies of the same gene. The affinity tag can be inserted at the N-terminus or inside the protein or, preferably, at the C-terminus in order to avoid interfering with the expression from the natural promoter. Two different tags can be attached to two different subunits of the protein complex. Two different tags can be attached to both ends of the protein. One tag can be attached to one end of the protein and another different tag can be inserted inside the gene. The binding domains of the affinity tag can be fused to the protein of interest directly or a by a linker sequence. The fusion protein can contain the whole protein of interest or only a part of it. It is best if the affinity tag is fused in such a way that it is accessible (i.e. can bind easily to the respective ligand) and does not affect the expression and folding of the protein and the

assembly of a natural protein complex. In addition, the affinity tag should not interfere with the protein interactions and should be placed outside the interface. Since different organisms prefer different codons for the same amino acid, the sequences of the affinity tags can be modified to ensure good expression.

[0092] When the protein of interest is present in the cell in low copy number and an additional purification step is needed before the separation of the interacting proteins from the fusion protein, a combination of two different affinity tags can be used. It is best if the second immobilization is performed by the stronger tag, i.e. the dissociation constant (K_d) for the binding tag:ligand is smaller, in order to avoid leakage of the affinity tagged protein during the elution. Separation from the first solid support can be performed by proteolytic cleavage or by a treatment with a chemical agent.

[0093] Purification of interacting proteins by fusing two affinity tags to the protein or subunit of the protein complex of interest can be described as follows: (a) binding the fusion protein by the first tag to the first ligand, (b) washing away the unbound proteins, (c) separating the fusion protein from the first ligand, (d) binding the fusion protein to the second ligand, (e) washing away the residual unbound proteins, (f) separating the interacting proteins from the fusion protein, which remains bound to the second ligand. The method of the invention can be carried out with any of the following combinations of first tag/second tag: GST/SpA, GST/SpG, GST/polyHis, MBP/SpA, MBP/SpG, MBP/polyHis, SBP/SpA, SBP/SpG. An affinity tag that remains bound to the respective ligand at ionic strength equivalent to 0.4-0.5M KCl is suitable for carrying out the invention when the elution is performed by increasing the ionic strength. Affinity tags genetically derived from Streptococcal protein G (SpG) or from Protein A from Staphylococcus aureus (SpA) bind strongly to IgG-sepharose and are suitable affinity tags for carrying out the invention. GST:glutathione binding is not affected by ionic strength equivalent to 0.5 M KCl and that makes the GST-tag a good candidate for a single tag for carrying out the invention or a second tag when a dual affinity tag is used. GST-tag is suitable for a first tag, too.
On the other hand, an affinity tag that can be separated from the respective ligand under conditions that do not disrupt the transient protein interactions can be used as a first tag. For example, GST-tag can be separated from glutathione beads by treatment with 5-20 mM reduced

glutathione. Flag-tag can be separated by adding Flag-tag (DYKDDDDK) and can be used as a first tag.
MBP can be separated by treatment with 10mM maltose and is suitable for a first tag.
MBP:maltose binding is not affected by ionic strength that is equivalent to 0.7-1 M KCl and is suitable for a second tag or as a single tag.
Poly-His tag requires stringent conditions for release of the immobilized protein and is not suitable for a first tag. It can be used as a second tag or as a single tag because the tag remains bound to Ni2+ column at 0.5M KCl.
The Myc tag (AEEQKLISEEDLLRK) and hemagglutinin tag (HA) (YPYDVPDYASL) are suitable for a first tag. It remains to be determined if they are good candidates for a second tag (i.e. if they remain bound to the ligand coated solid support at ionic strength that is equivalent to 0.4 or 0.5 M KCl) or for a single tag. The HA-tag is more suitable for a second tag or single tag because it contains only two electrostatically charged amino acids.
Other tags that can be used for carrying out the invention are the T7-Tag sequence (the initial 11aa of the T7 gene), VSV-G (15 AA C-terminal peptide of VSV-G), b-GAL, Thioredoxin. The list illustrates but does not limit the scope of the invention. The review of K. Terpe in Appl. Microbiol. Biotechnol. 2003) describes the most popular affinity tags.

[0094] The invention can be carried out with whole cells, cellular organelles, cellular lysate, or other biological fluids. In addition, the invention can be performed with fusion protein and interacting proteins expressed in a cell free system. Cellular lysate can be fractionated in order to obtain a fraction that is enriched in the protein of interest. When the organism is a multicellular organism, the immobilized fusion protein can originate from one tissue but the interacting proteins and/or other biomolecules can originate from a different tissue (e.g. hormones). In this case, the biological fluid containing the interacting biomolecules can be applied to the fusion protein before or after the immobilization. When the biological fluid containing the interacting biomolecules is applied after the immobilization of the fusion protein and removal of the unbound substances, the immobilized fusion protein can be associated with biomolecules from its original tissue or, the associated biomolecules can be removed before applying the other biological fluid originating from a different tissue. The fusion protein can be cross-linked to the affinity matrix in order to minimize the leakage.

[0095] When the protein of interest is present in low copy number in the cell, it can be expressed from a different promoter that generates more copies of the protein. Optionally, the promoter can be inducible and if the protein of interest is a subunit of a stable protein complex, the other subunits and the fusion protein of interest can be expressed from identical inducible promoters in order to ensure proper stoichiometries of the stable protein complex. In this case, great care should be taken not to over-express the protein of interest to a level that leads to formation of inclusion bodies.

[0096] The fusion protein can originate from one organism and the interacting proteins can originate from a different organism (e.g. host organism and parasite organism). The affinity tagged protein of interest can be exogenous or endogenous. However, the invention gives the most valid results when the fusion protein and the interacting proteins are expressed in their endogenous organism and their expression is driven from their natural promoters.

[0097] In most cases, the affinity tag binds to the ligand directly, but in some cases, one or more bridging molecules can mediate the binding. Indirect binding of an interacting protein to the fusion protein can be mediated by another protein and/or by a nucleic acid and/or by other biomolecule.

[0098] Figure 20 illustrates the problems with analyzing fractions or gel bands containing proteins with different abundances by mass spectrometry. After the trypsin digest, the concentration range of the peptides remains approximately the same and it results in mass spectrum containing peaks with different intensity. Left panel: When the mass spectrum contains large and small peaks separated by only a few (1-5) mass units, the detection of the latter is impossible because of the dynamic range problem in mass spectrometry. Analyzing such spectrum leads to the identification of only the major protein(s) in the band. Right panel: High background noise problem in mass spectrometry - when a protein mixture with high dynamic range (ratio between the most abundant and the least abundant protein) is analyzed directly the small peptides are not detected because their intensities are lower than the background. Legend: 450 – high abundance peptide; 451 – different low abundance peptides; 452 – background of the mass spectrum; 453 – low abundance peptide is not detected because it is obscured by a high abundance peptide which mass is two mass units smaller; 454 – the separation of the substoichiometrically interacting proteins from the

high abundance protein of interest results in lower background (due to the absence of the high abundance peptides) and makes possible the detection of the low abundance peptides.

Claims:

What is claimed in this application is:

1. Method for isolating biomolecules that associate in vivo with a biomolecule of interest comprising the steps: (a) obtaining a cell or cellular lysate or fraction of cellular lysate or other biological fluid containing naturally assembled biological complexes that include the biomolecule of interest and the associated biomolecules, (b) immobilizing selectively the biomolecule of interest on an affinity matrix and removing the unbound substances, (c) isolating and/or analyzing one or more associated biomolecules by separating them from the immobilized biomolecule of interest which remains bound to the affinity matrix during the separation.

2. Method for isolating and/or analyzing proteins and/or other substances that associate in vivo with a polypeptide of interest that is fused to at least one affinity tag, comprising the steps: (a) introducing in an organism or cell line a heterologous nucleic acid encoding a polypeptide of interest or part of a polypeptide of interest fused to an affinity tag that can bind selectively to an affinity matrix, (b) expressing the fusion protein under physiological conditions that enable the formation of stable and transient protein complexes, (c) immobilizing the fusion protein via the affinity tag on the affinity matrix and removing proteins and/or other substances that are not bound directly or indirectly to the fusion protein, (c) isolating and/or analyzing one or more associated proteins and/or other substances by separating them from the immobilized fusion protein which remains bound to the affinity matrix during the separation.

3. Method for isolating and/or analyzing proteins and/or other substances that associate in vivo with a polypeptide of interest that is fused to at least two different affinity tags, comprising the steps:
(a) introducing in an organism or cell line a heterologous nucleic acid encoding a polypeptide of interest or part of a polypeptide of interest fused to at least two different affinity tags that can bind selectively to different affinity matrixes, (b) expressing the fusion protein under physiological conditions that enable formation of stable and transient protein complexes, (c) immobilizing the fusion protein via the first affinity tag on a first affinity matrix and removing proteins and/or other substances that are not bound directly or indirectly to the fusion protein, (d) separating the fusion

protein from the first affinity matrix and immobilizing it via the second affinity tag on the second affinity matrix and removing proteins and/or substances that are not bound directly or indirectly to the fusion protein, (e) isolating and/or analyzing one or more associated proteins and/or other substances by separating them from the immobilized fusion protein which remains bound to the affinity matrix during the separation.

4. Method for isolating and/or analyzing proteins and/or other substances that associate in vivo with a protein complex, two or more subunits of which are fused to different affinity tags, comprising the steps: (a) introducing in an organism or cell line two different heterologous nucleic acids encoding two different polypeptides of interest or part of a polypeptides of interest, fused to different affinity tags that can bind selectively to different affinity matrixes, and said polypeptides being part of a protein complex, (b) expressing the fusion proteins in the organism under physiological conditions that enable formation of stable and transient protein complexes, (c) immobilizing the protein complex via the first affinity tag on the first affinity matrix and removing proteins and/or substances that are not bound directly or indirectly to the complex, (d) separating the complex from the first affinity matrix and immobilizing it via the second affinity tag on the second affinity matrix and removing proteins and/or substances that are not bound directly or indirectly to the complex, (e) isolating and/or analyzing one or more associated proteins and/or other substances by separating them from the immobilized fusion protein which remains bound to the affinity matrix during the separation.

5. Method according to claims 1 or 2, wherein the protein of interest or other biomolecule is immobilized selectively on an affinity matrix coated with antibody that has been raised against the protein of interest.

6. Method according to claim 2, wherein one or more affinity tags can bind selectively to the Fc domains of immunoglobulin.

7. Method according to claim 6, wherein one or more affinity tags contain one or more IgG binding regions of Staphylococcus aureus Protein A (SpA-tag) or Streptococcal protein G (SpG-tag).

8. Method according to claim 2, wherein one or more affinity tags can be separated selectively from the affinity matrix by treatment with a chemical agent.

9. Method according to claim 8, wherein the affinity tag is GST-tag or SBP-tag or a calmodulin binding peptide or maltose binding peptide.

10. Method according to claims 3 or 4, wherein the fusion protein is separated from the first affinity matrix by enzymatic cleavage.

11. Method according to claim 10, wherein the enzymatic cleavage is cleavage by TEV protease.

12. Method according to claims 3 or 4, wherein the first immobilization is performed by binding to a solid support coated with a specific antibody and the first release is performed by addition of the same antibody.

13. Method according to claims 3 or 4, wherein the second immobilization is performed by binding to a solid support coated with specific antibody.

14. Method according to claims 3 or 4, wherein the first tag can bind selectively to the Fc domains of immunoglobulin.

15. Method according to claim 14, wherein the first tag contains one or more IgG binding regions of Staphylococcus aureus Protein A (SpA-tag) or Streptococcal protein G (SpG-tag).

16. Method according to claims 3 or 4, wherein the first affinity tag can be separated selectively from the first affinity matrix by treatment with a chemical agent.

17. Method according to claim 16, wherein the first affinity tag is GST-tag or SBP-tag or a calmodulin binding peptide or maltose binding peptide.

18. Method according to claims 3 or 4, wherein the second affinity tag can be separated selectively from the first affinity matrix by treatment with a chemical agent.

19. Method according to claim 18, wherein the second affinity tag is GST-tag or SBP-tag or a calmodulin binding peptide or maltose binding peptide.

20. Method according to claims 3 or 4, wherein the second affinity tag can bind selectively to the Fc domains of immunoglobulin.

21. Method according to claim 20, wherein the second tag contains one or more IgG binding regions of Staphylococcus aureus Protein A (SpA-tag) or Streptococcal protein G (SpG-tag).

22. Method according to one of the previous claims, wherein the associated proteins are separated from the immobilized protein of interest by changing the concentration of at least one chemical agent.

23. Method according to claim 22, wherein the chemical agent is KCl.

24. Method according to claim 22, wherein the change of the concentration of the chemical agent is between 0 mM and 30 mM.

25. Method according to claim 22, wherein the change of the concentration of the chemical agent is between 30 mM and 300 mM.

26. Method according to claim 22, wherein the change of the concentration of the chemical agent is between 300 mM and 700 mM.

27. Method according to claim 22, wherein the change of the concentration of the chemical agent is between 700 mM and 2 M.

28. Method according to one of the previous claims, wherein the associated proteins are separated from the immobilized fusion by changing the pH of the medium.

29. Method according to one of the previous claims, wherein the associated proteins are separated from the immobilized fusion protein by enzymatic treatment that modifies the fusion protein.

30. Method according to one of the previous claims, wherein the associated proteins are separated from the immobilized fusion protein by enzymatic treatment that modifies the associated proteins.

31. Method for identification and elimination of the contaminant proteins (proteins that bind nonspecifically to the ligand and/ or the affinity matrix and/or to any protein) in an organism or cell line comprising preparation of cellular extract or other biological fluid from the said organism or cell line devoid of an affinity-tagged gene and performing all the purification steps in the same way as with organism containing affinity tagged protein.

32. Method for identification and elimination of the contaminant proteins (proteins that bind nonspecifically to the ligand and/ or the affinity matrix and/or to any protein) in an organism or cell line comprising comparing the proteins that bind to several biologically unrelated fusion proteins from the said organism or cell line and identifying the common ones as contaminants.

33. Method according to one of the previous claims, wherein after the separation of the associated proteins and/or other biomolecules from the fusion protein, the immobilized protein or a protein complex is mixed again with protein extract or other biological fluid from the same or different organism and, after removal of unbound substances, the associated proteins are separated from the immobilized fusion protein.

34. Method according to claim 33, wherein an affinity tagged protein is covalently cross-linked to the ligand coated affinity matrix after the first elution.

35. Method according to one of the previous claims, wherein a chemical or a biomolecule is identified as a drug or pre-drug by its capability to affect selectively the separation of the associated proteins and/or other biomolecules from the protein of interest when it is added to or removed from the cells and/or cellular lysate and/or other biological fluid and/or buffers.

36. Method according to claim 35, wherein the fusion protein contains at least one mutation.

37. Method according to claim 35, wherein the fusion protein associates directly or indirectly with another protein that contains at least one mutation.

38. Method according to claims 35, 36 and 37, wherein the chemical or biomolecule is designed and/or synthesized and/or selected for testing by the following features: (a) capability to bind selectively to the protein target, and (b) containing at least one electrostatic charge that is identical to the charge that has been changed as a result of the mutation, and (c) after the binding of the chemical to the protein target, the electrostatic charge is located at distance between 0 and 0.5 nanometer from the mutated amino acid.

39. Method according to one of the previous claims, wherein the protein of interest is an enzyme.

40. Method according to one of the previous claims, wherein the protein of interest is a substrate for an enzyme.

41. Method according to one of the previous claims, wherein the protein of interest is an enzyme and a substrate for another enzyme.

42. Method according to claim 41, wherein the protein of interest is RNA Polymerase or DNA Polymerase.

43. Method for identifying protein-protein association as a putative cause for a disease, comprising associating the disease with a mutation that changes the electrostatic properties of the protein by replacing Lysine and/or Arginine and/or Aspartate and/or Glutamate and/or Histidine with an uncharged or oppositely charged amino acid.

44. Method for identifying protein-protein association as a putative cause for a disease, comprising associating the disease with a mutation that changes the electrostatic properties of the protein by replacing an amino acid with Lysine and/or Arginine and/or Aspartate and/or Glutamate and/or Histidine.

45. Method according to claim 1, wherein the biomolecule of interest is a glycoprotein and the affinity matrix is lectin coated beads.

46. Method according to claim 1, wherein the biomolecule of interest is a nucleic acid, which is part of a nucleoprotein complex and the affinity matrix, consists of immobilized nucleic acid with complementary sequence.

47. Method according to claim 1, wherein the biomolecule of interest is a nucleic acid, which is part of a nucleoprotein complex and is genetically engineered so that it contains a poly-Guanosine and the affinity matrix consists of immobilized poly-dCytosine.

48. Reagent kit comprising at least one chemical agent for separating the associated proteins from the protein of interest.

49. Reagent kit comprising a buffer for preparation of protein extract and washing buffer.

Figure 1

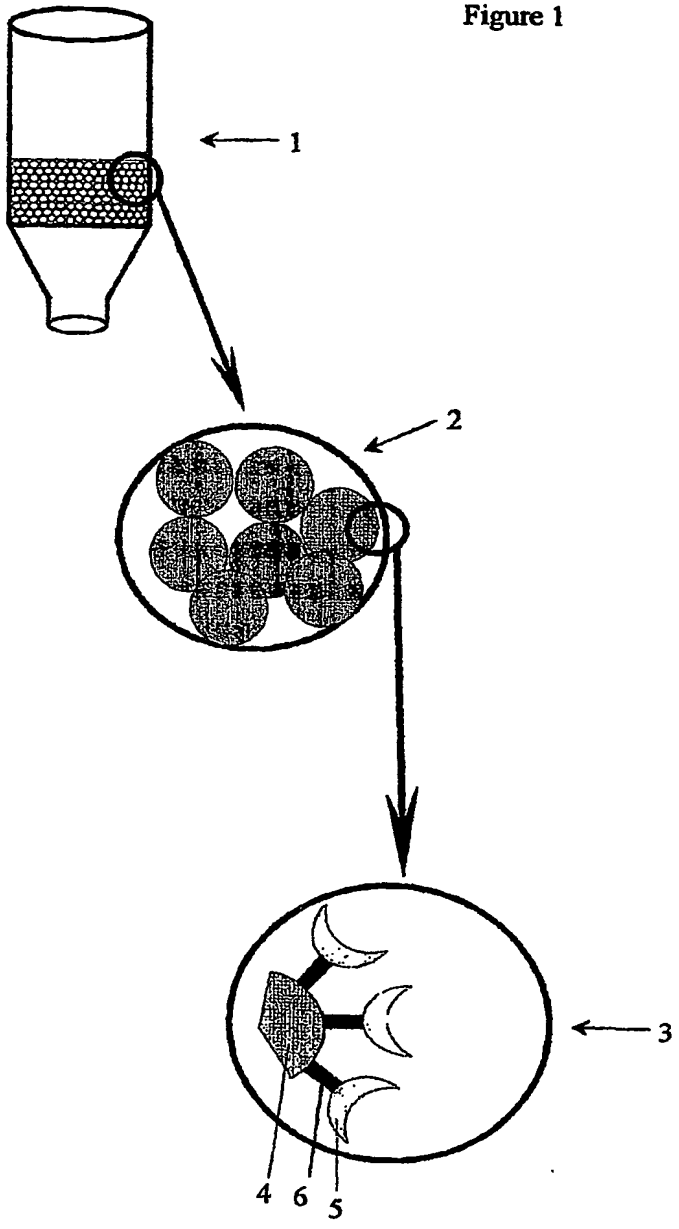


Figure 2

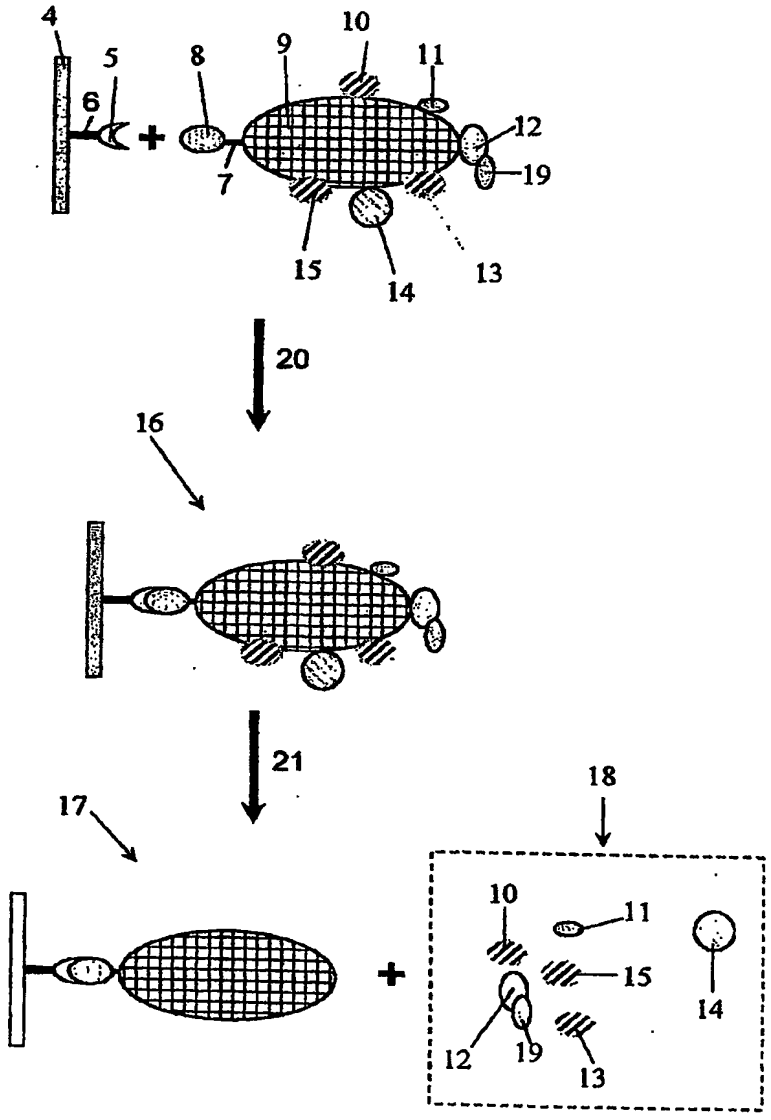


Figure 3

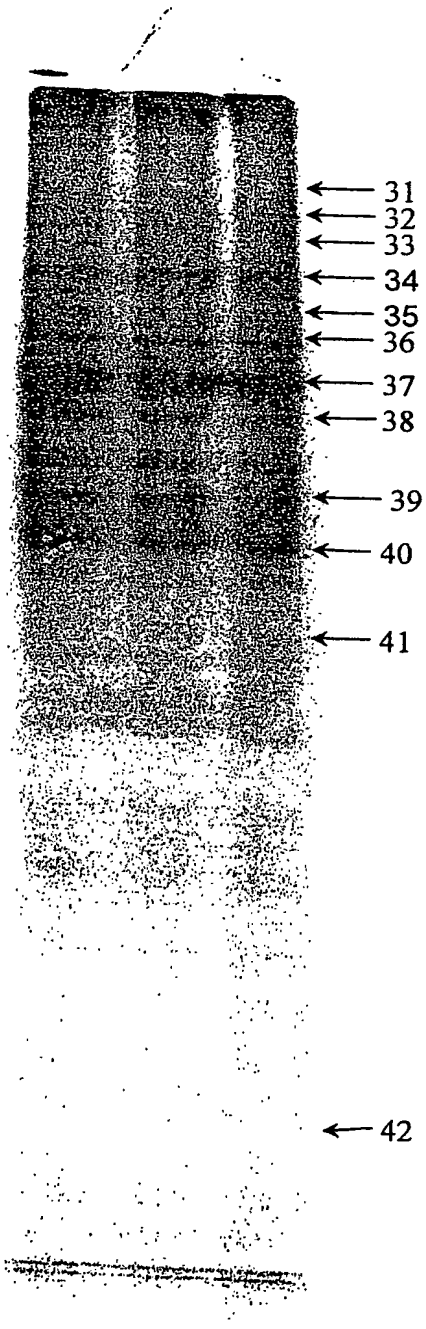


Figure 4

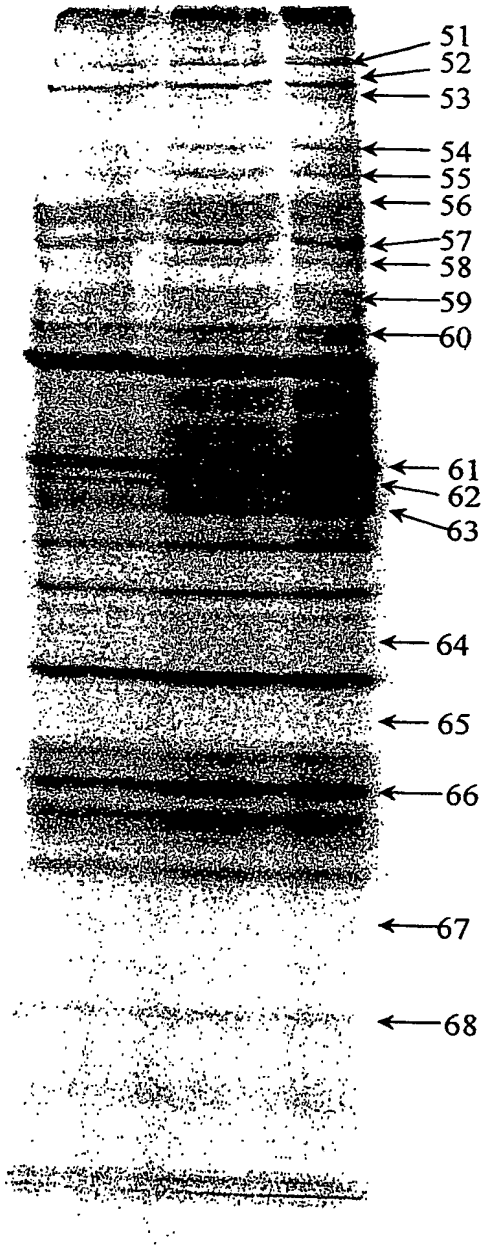


Figure 5

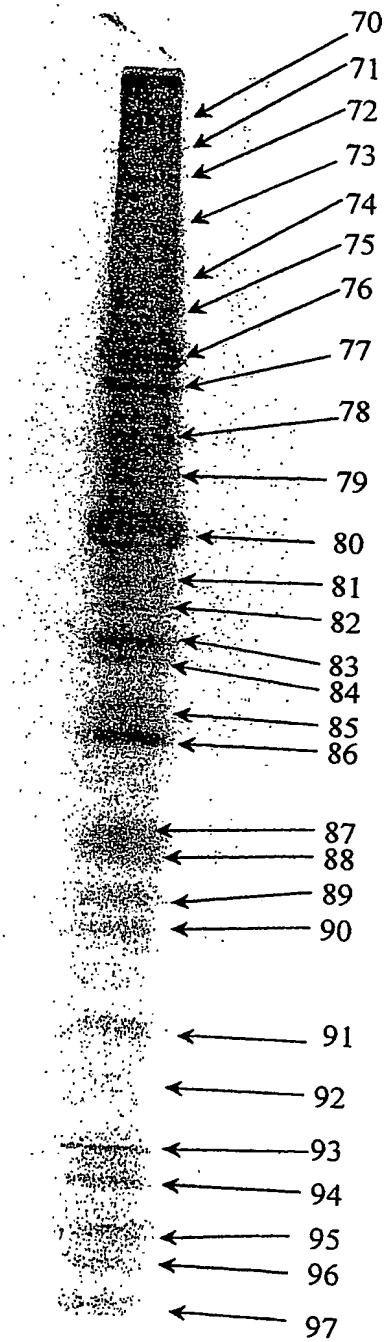

70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97

Figure 6

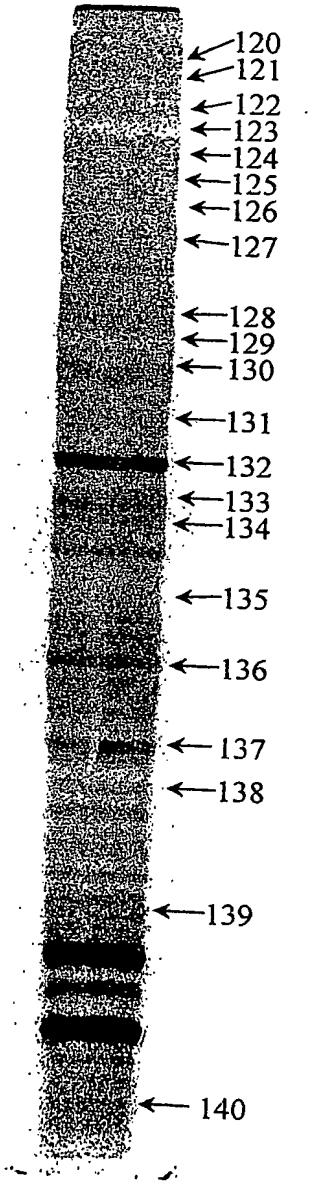

120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140

Figure 7

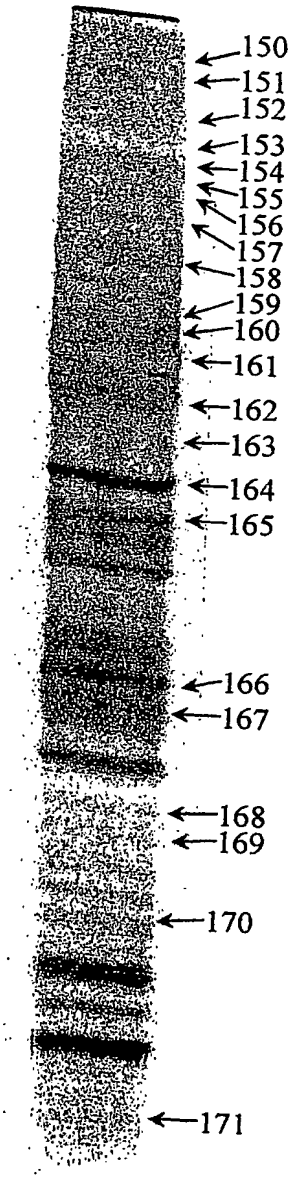

150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171

Figure 8

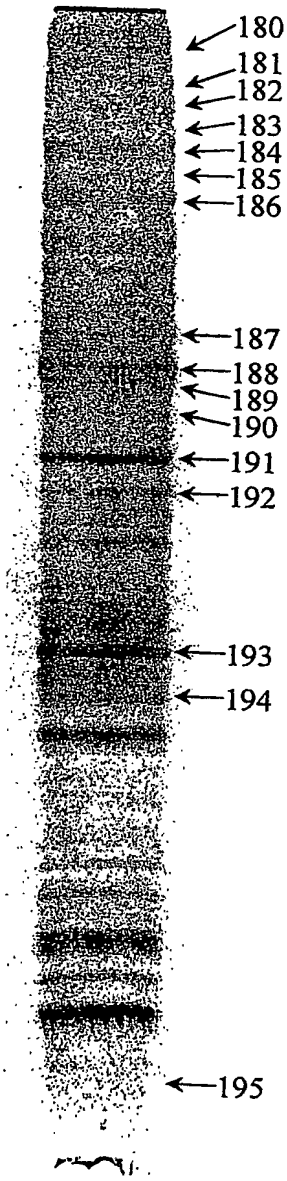


Figure 9

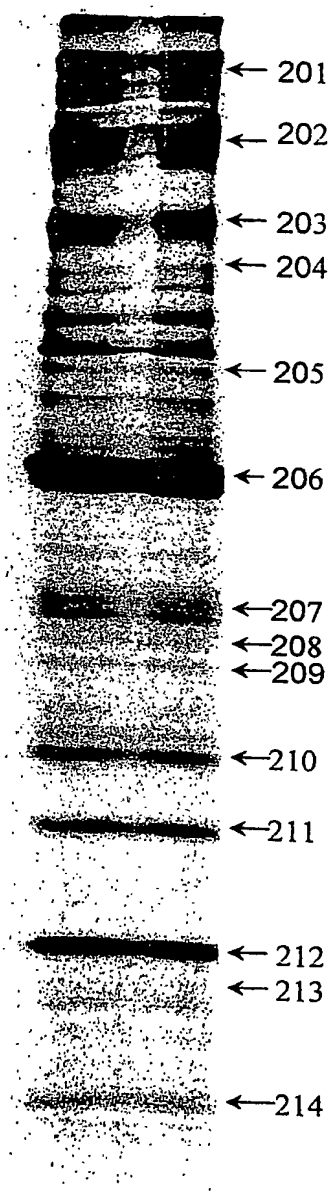


Figure 10

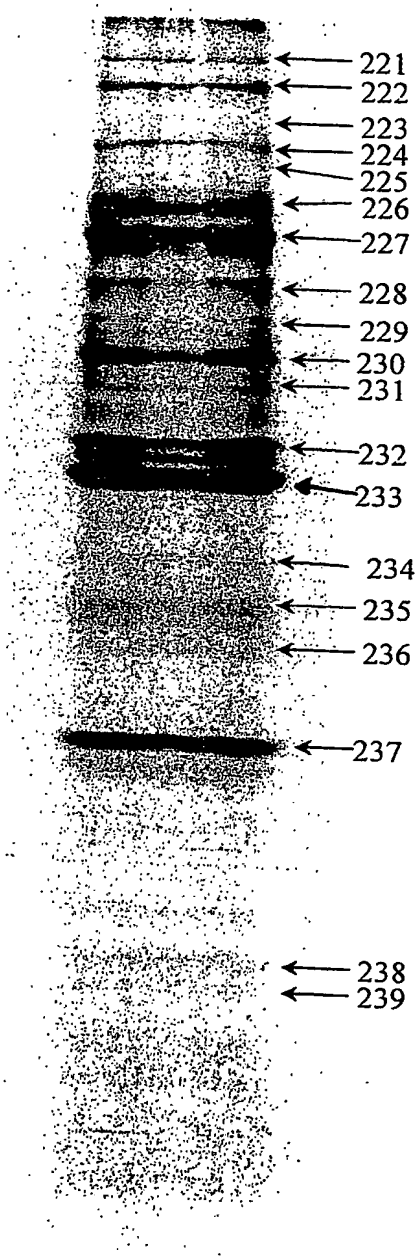

221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239

Figure 11

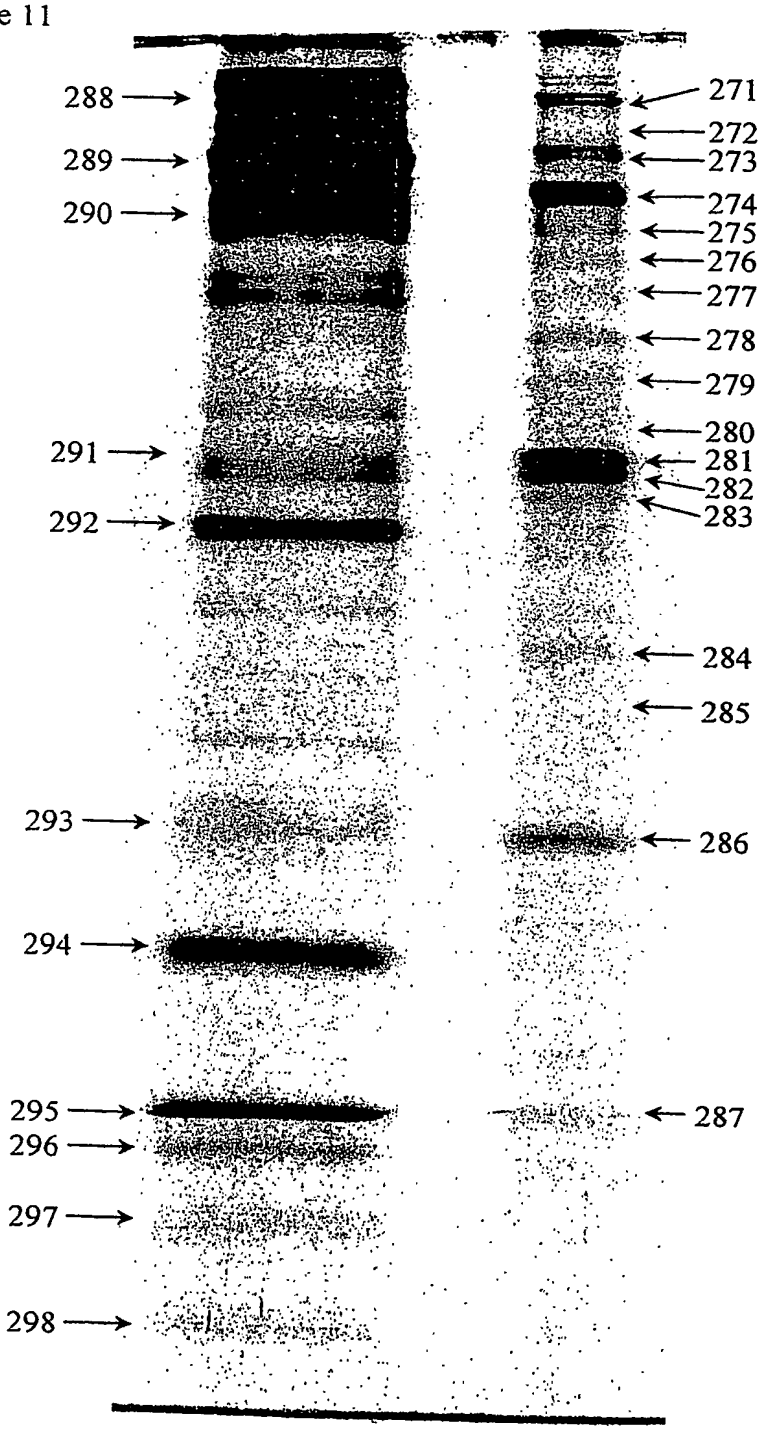

288
289
290
291
292
293
294
295
296
297
298
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287

Figure 12

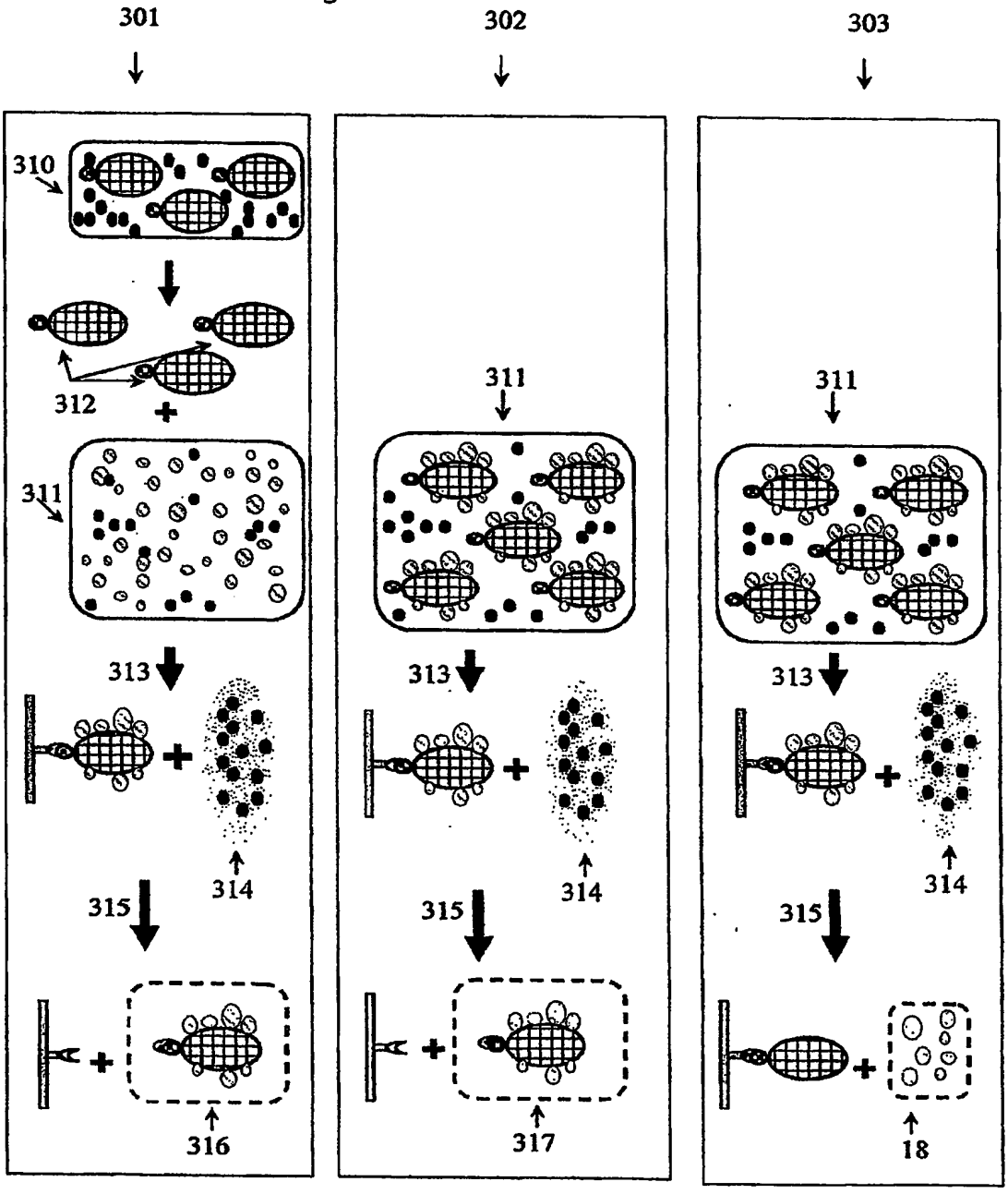

301
302
303
310
312
311
313
314
315
316
311
313
314
315
317
311
313
314
315
18

Figure 13

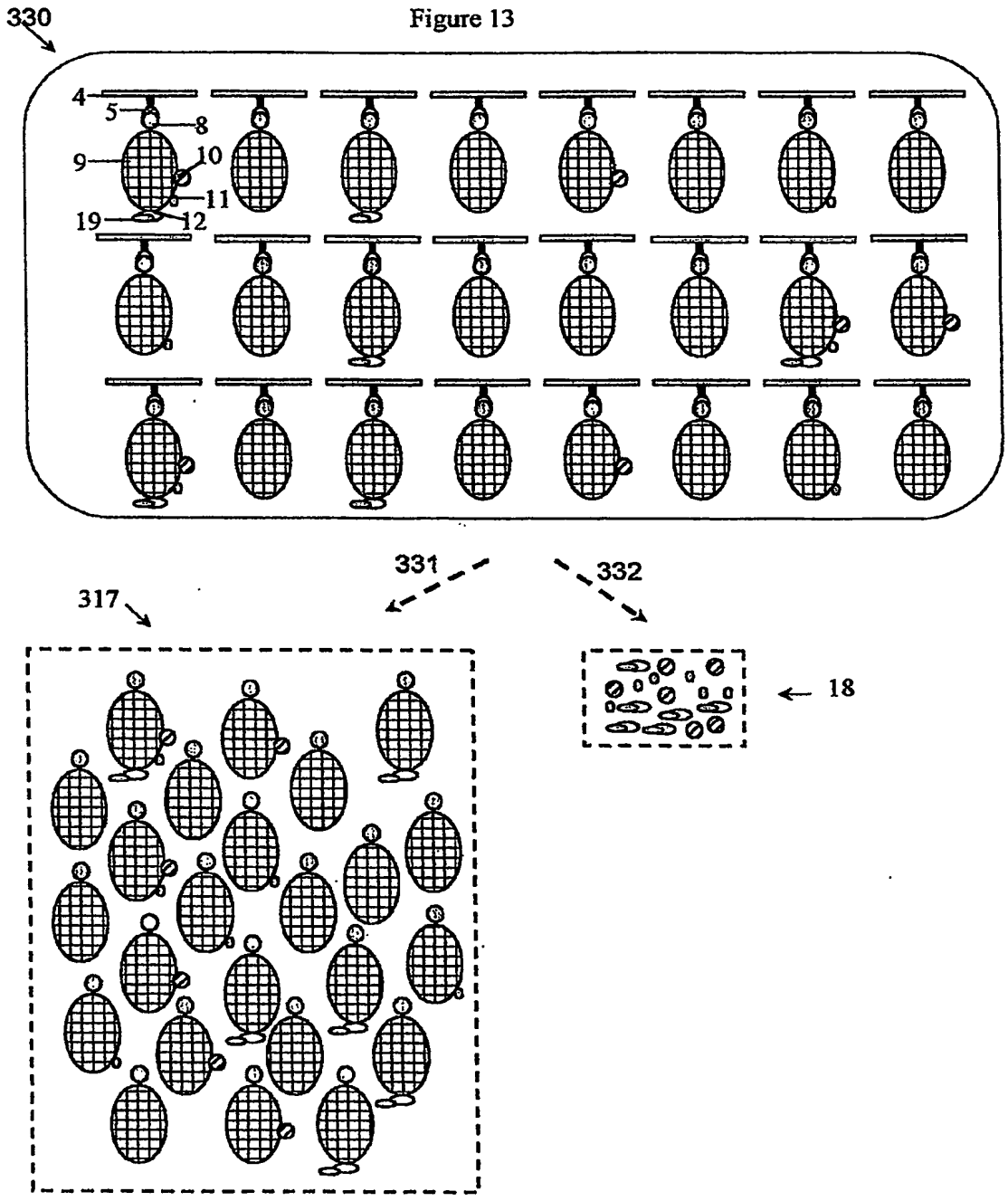


Figure 14

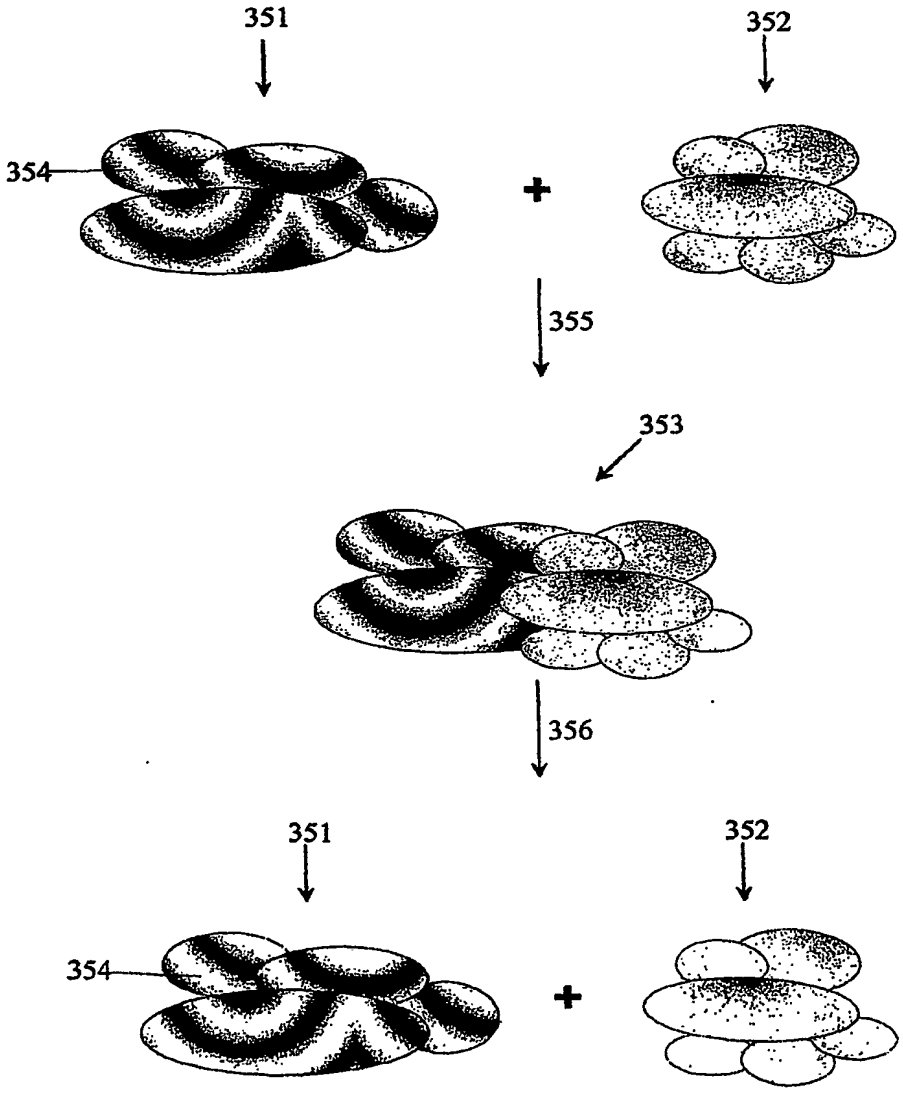


Figure 15

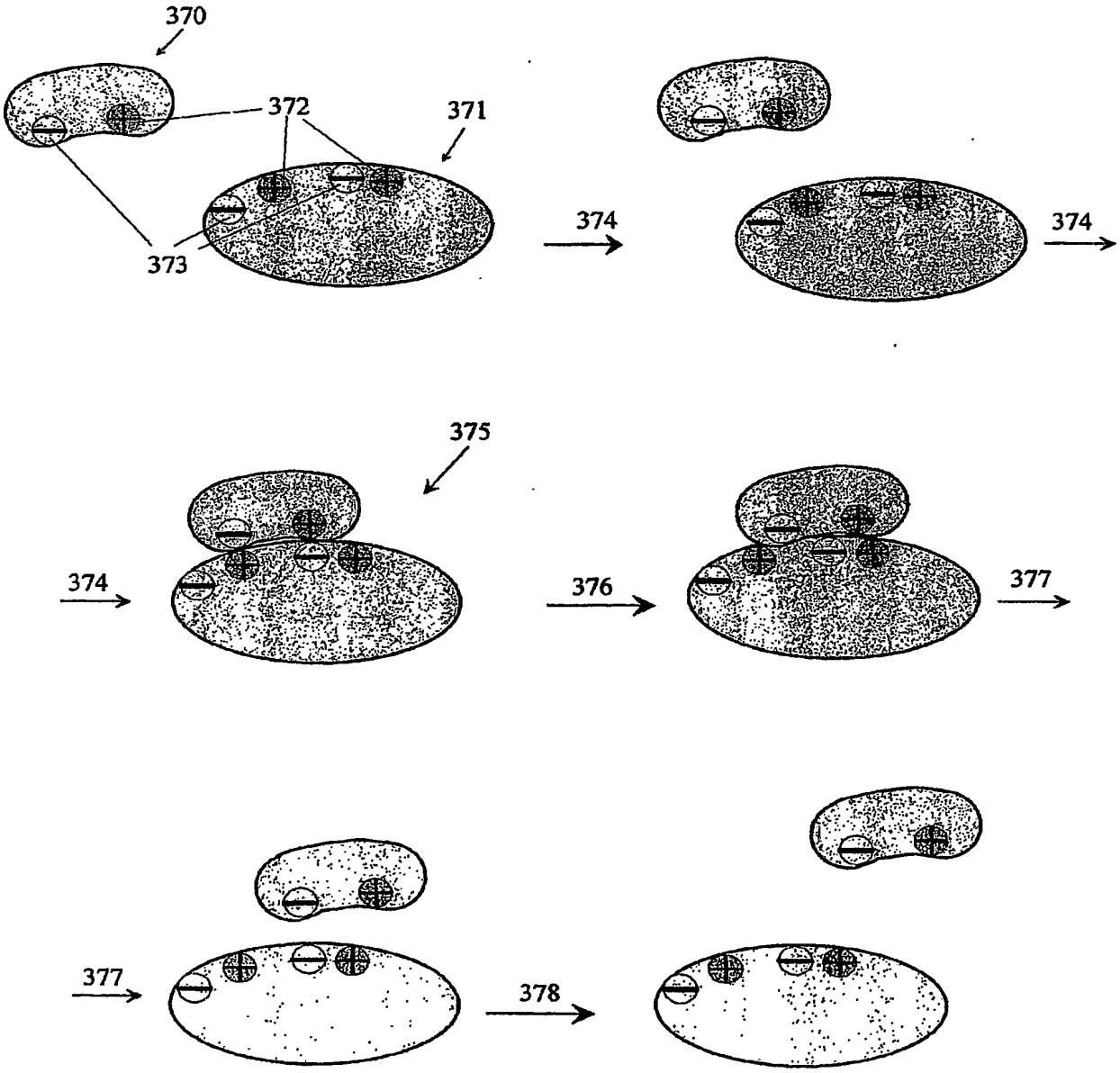
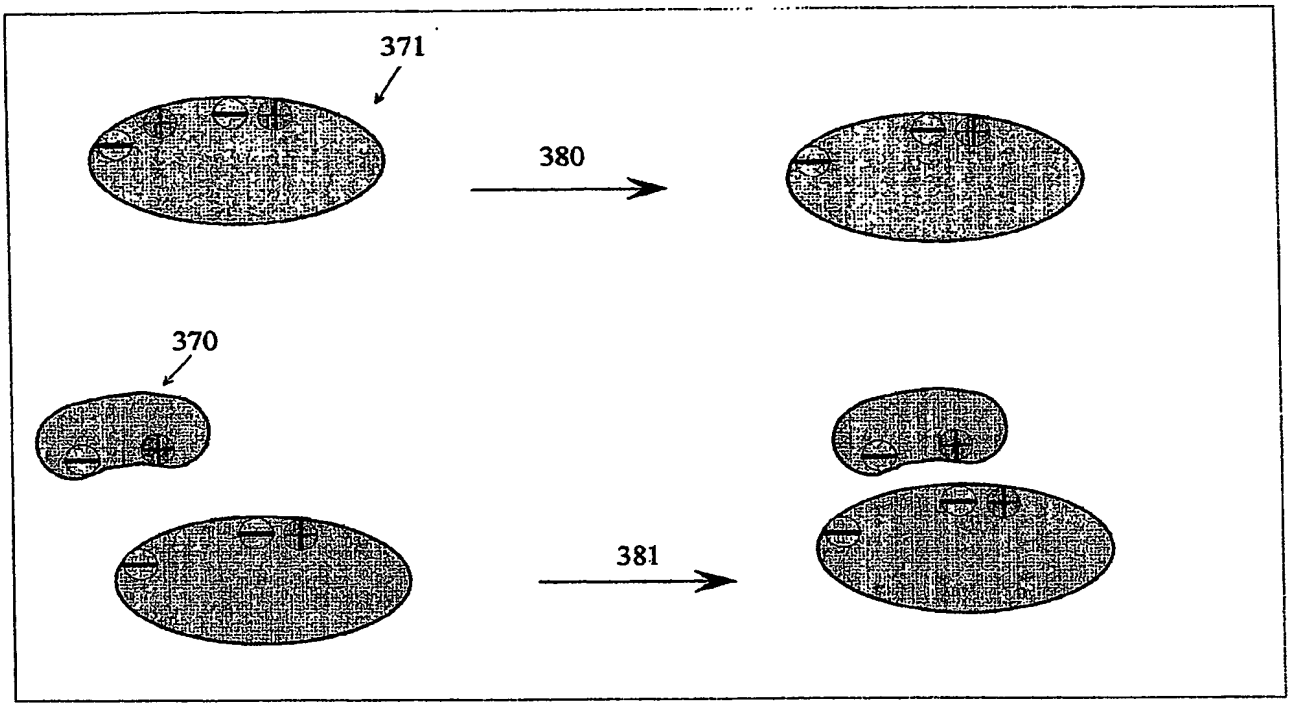
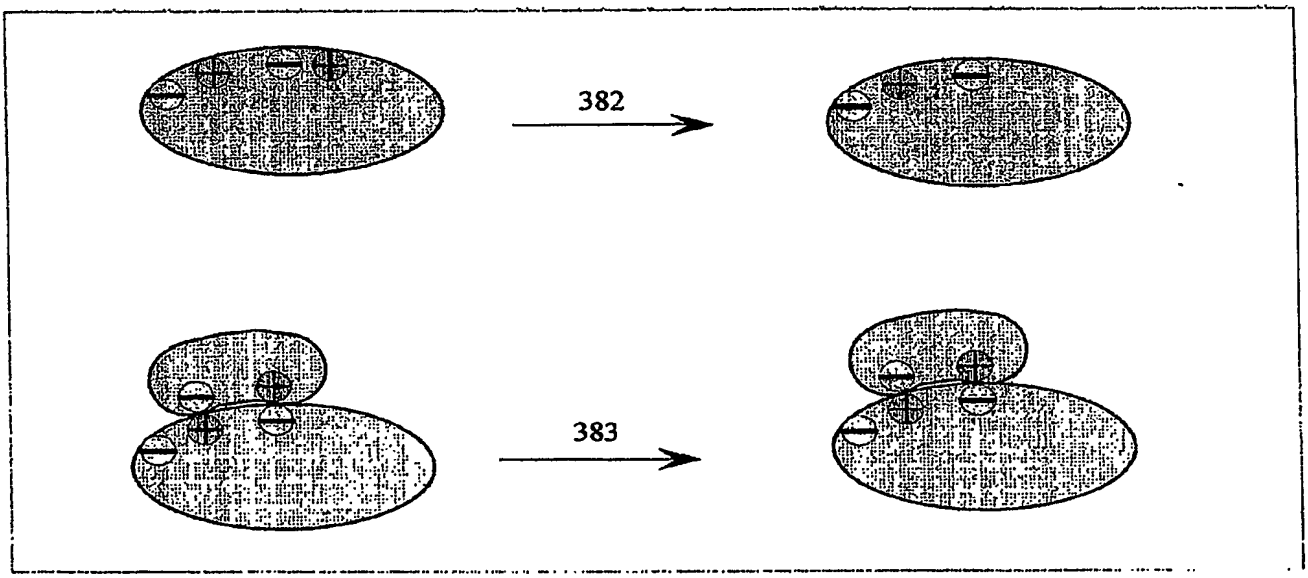


Figure 16





17/21

Figure 17

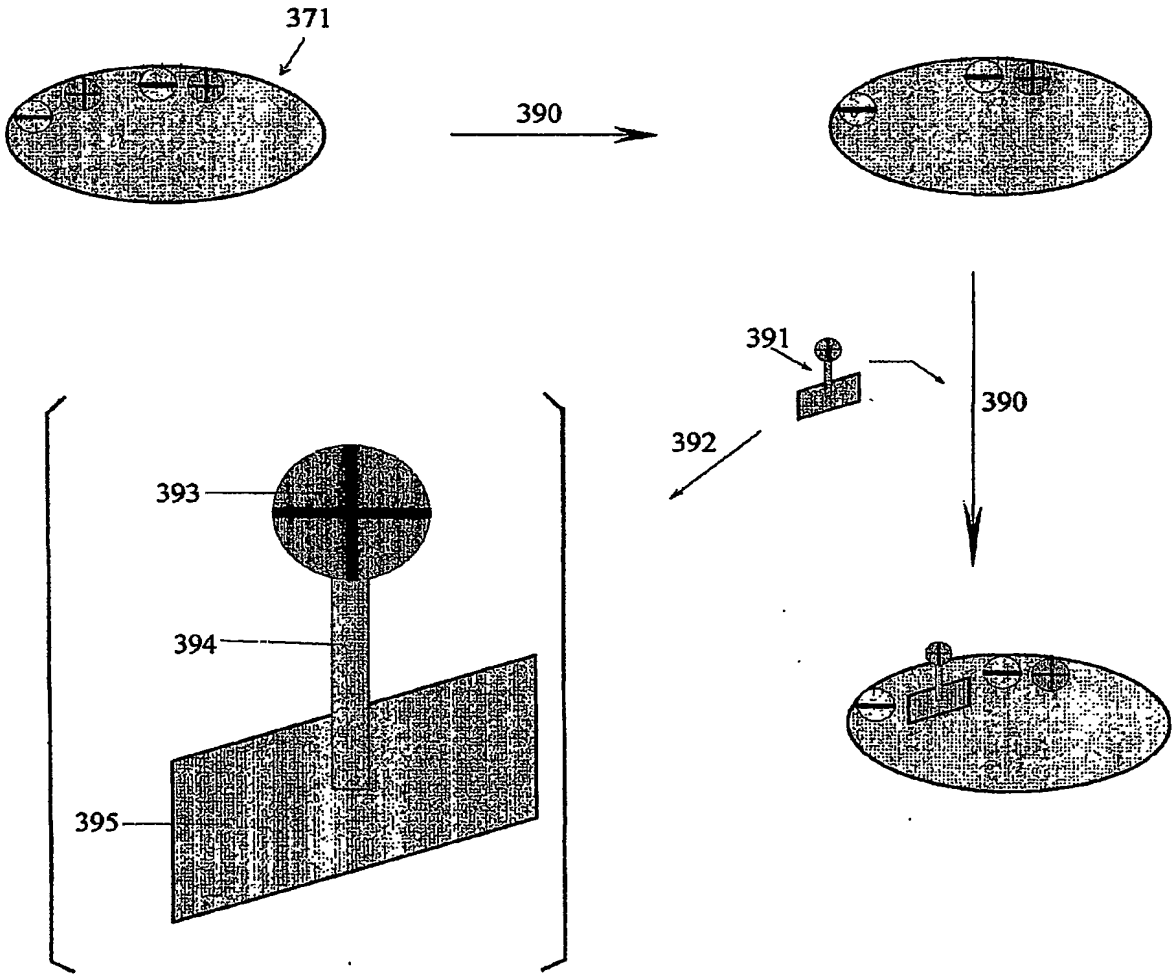


Figure 18

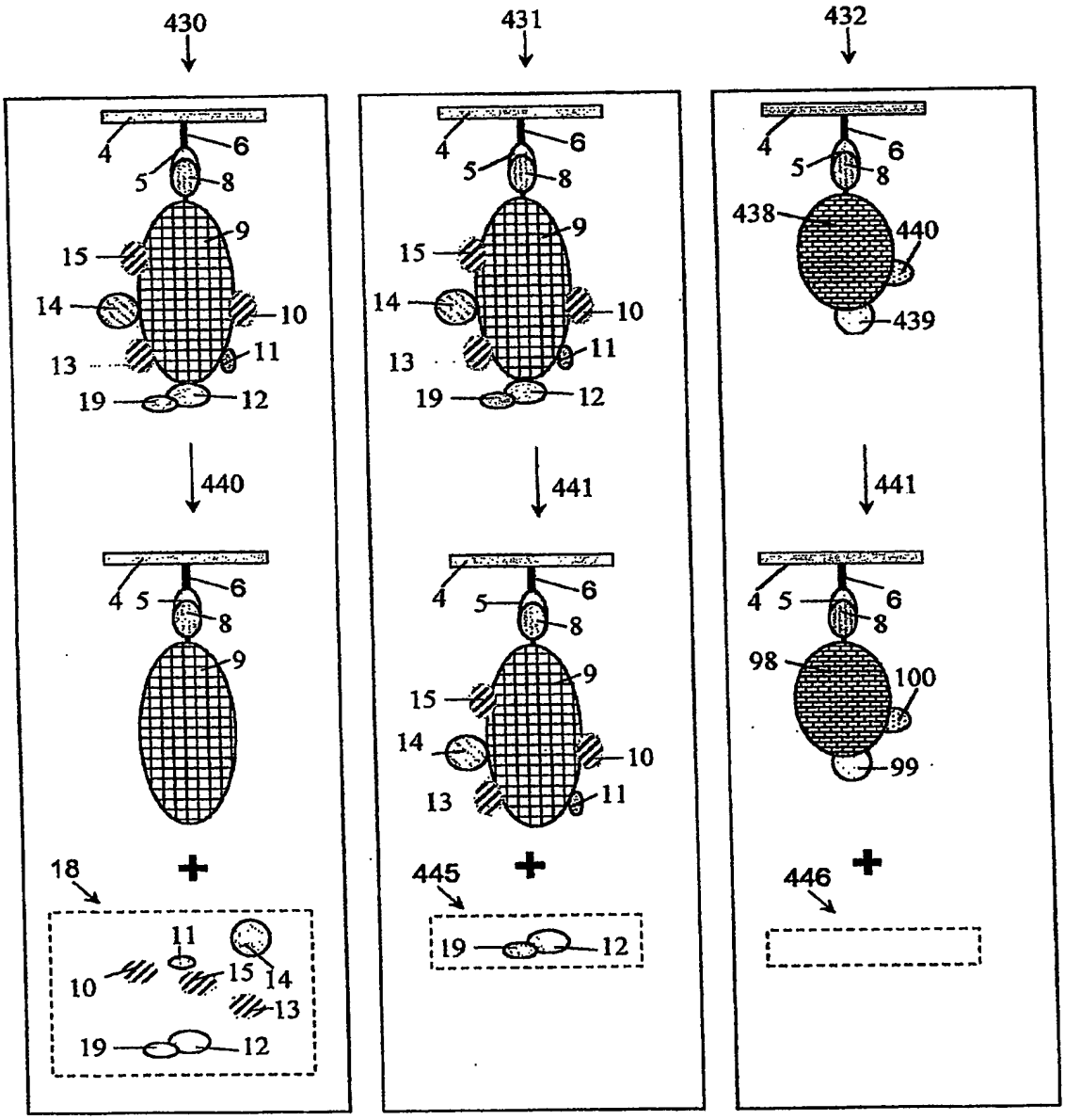

430
431
432
4
5
6
8
9
15
14
10
13
11
19
12
438
440
439
440
441
441
98
100
99
18
445
446
11
10
15
14
13
19
12

Figure 19

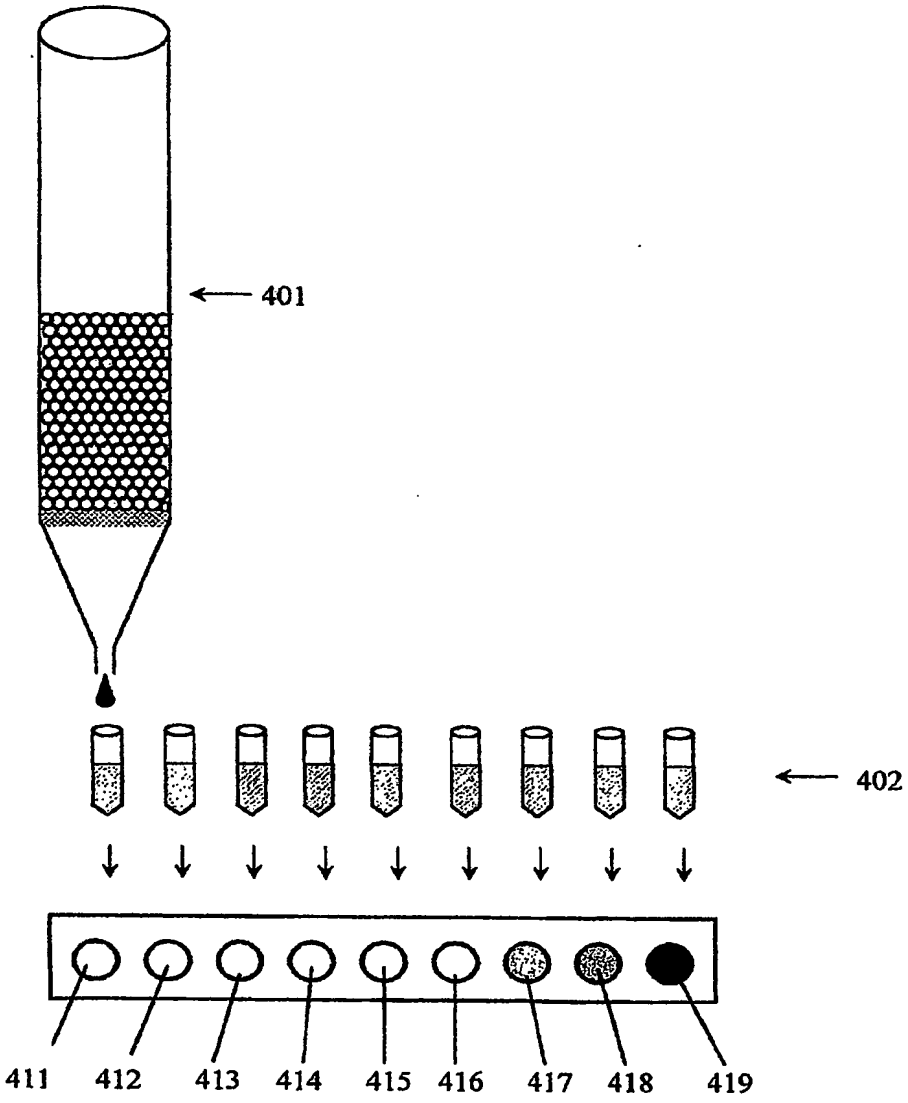


Figure 20

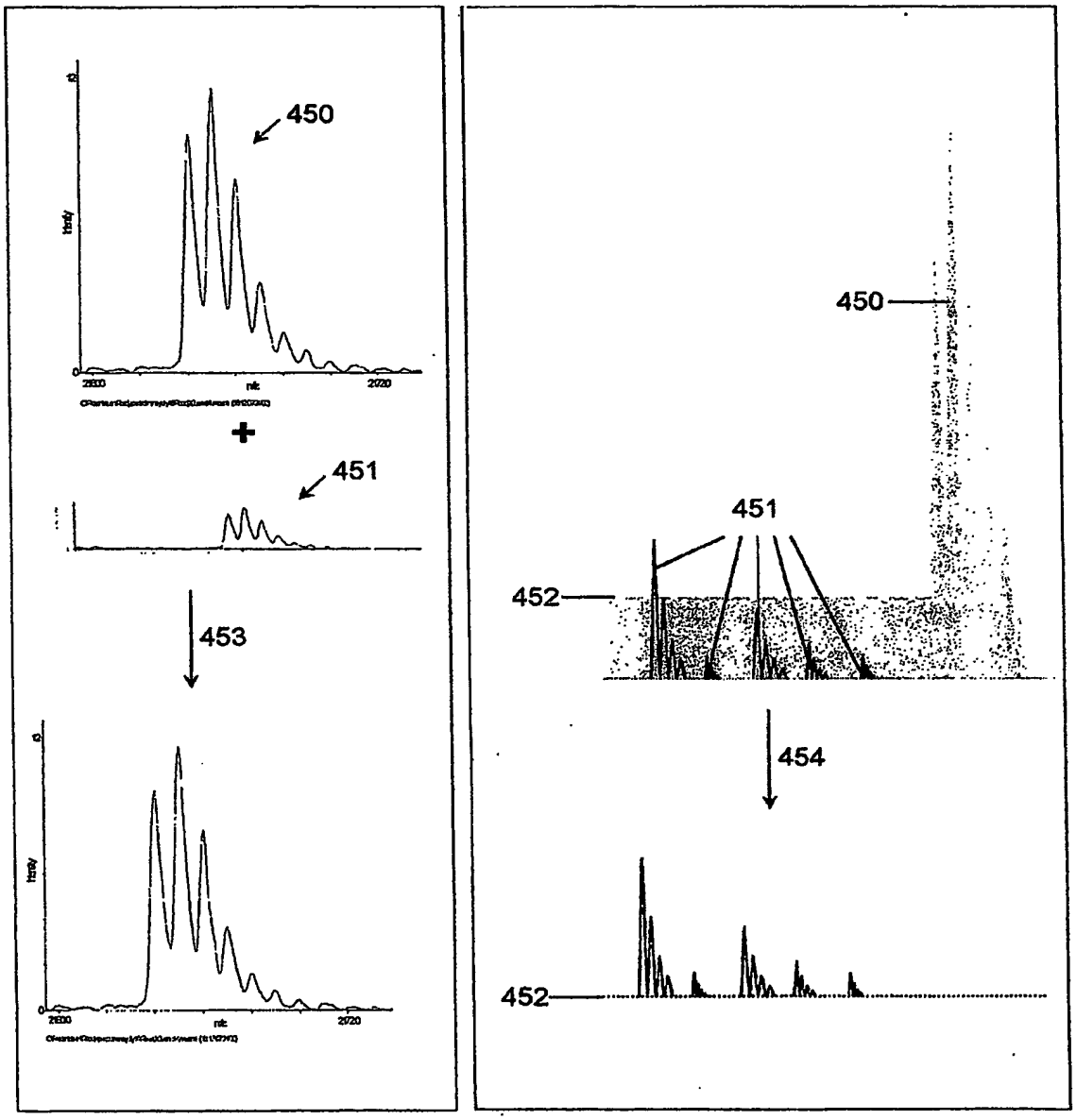


Figure 21